



**UNIVERSIDAD TÉCNICA DE AMBATO**

**FACULTAD DE INGENIERÍA EN SISTEMAS ELECTRÓNICA E  
INDUSTRIAL**

**CARRERA DE INGENIERÍA EN ELECTRÓNICA Y  
COMUNICACIONES**

**Tema:**

---

**“SISTEMA DE CODIFICACIÓN DE SEÑALES DE VOZ A TEXTO USANDO  
MATLAB PARA PERSONAS CON DEFICIENCIA AUDITIVA DEL  
INSTITUTO DE EDUCACIÓN ESPECIAL DR. CAMILO GALLEGOS”**

---

Proyecto de Trabajo de Graduación. Modalidad: TEMI. Trabajo Estructurado de Manera Independiente, presentado previo la obtención del título de Ingeniero en Electrónica y Comunicaciones.

AUTOR: Jaime Andrés Cárdenas Córdova

TUTOR: Ing. Luis Alfredo Pomaquero Moreno

Ambato - Ecuador

Diciembre-2012

## **APROBACION DEL TUTOR**

En calidad de Tutor del Trabajo de Investigación sobre el tema:

“SISTEMA DE CODIFICACIÓN DE SEÑALES DE VOZ A TEXTO USANDO MATLAB PARA PERSONAS CON DEFICIENCIA AUDITIVA DEL INSTITUTO DE EDUCACIÓN ESPECIAL DR. CAMILO GALLEGOS”, de Jaime Andrés Cárdenas Córdova, estudiante de la Carrera de Ingeniería en Electrónica y Comunicaciones, de la Facultad de Ingeniería en Sistemas Electrónica e Industrial, de la Universidad Técnica de Ambato, considero que dicho informe investigativo reúne los requisitos y méritos suficientes para ser sometidos a la evaluación del Tribunal de Grado, que el Honorable Consejo Directivo designe, para su correspondiente estudio y calificación.

Ambato, 10 de Octubre 2012

EL TUTOR

-----  
Ing. Luis Alfredo Pomaquero Moreno

## **AUTORÍA**

El presente trabajo de investigación “SISTEMA DE CODIFICACIÓN DE SEÑALES DE VOZ A TEXTO USANDO MATLAB PARA PERSONAS CON DEFICIENCIA AUDITIVA DEL INSTITUTO DE EDUCACIÓN ESPECIAL DR. CAMILO GALLEGOS” es absolutamente original, auténtico y personal, en tal virtud, el contenido, efectos legales y académicos que se desprenden del mismo son de exclusiva responsabilidad del autor.

Ambato, 10 de Octubre de 2012

.....  
Jaime Andrés Cárdenas Córdova

C.C. 1804314506

## **APROBACIÓN DE LA COMISIÓN CALIFICADORA**

La Comisión Calificadora del presente trabajo conformada por los señores docentes, Ing. Edwin Rodrigo Morales Perrazo, e Ing. Geovanni Danilo Brito Moncayo, revisó y aprobó el Informe Final del trabajo de graduación titulado SISTEMA DE CODIFICACIÓN DE SEÑALES DE VOZ A TEXTO USANDO MATLAB PARA PERSONAS CON DEFICIENCIA AUDITIVA DEL INSTITUTO DE EDUCACIÓN ESPECIAL DR. CAMILO GALLEGOS, presentado por el señor Jaime Andrés Cárdenas Córdova de acuerdo al Art. 17 del Reglamento de Graduación para obtener el título Terminal de tercer nivel de la Universidad Técnica de Ambato.

Ambato Octubre 25, 2012

Ing. Oswaldo Eduardo Paredes Ochoa, M.Sc.

PRESIDENTE DEL TRIBUNAL

Ing. M.Sc. Edwin Rodrigo Morales Perrazo    Ing. M.Sc. Geovanni Danilo Brito Moncayo

DOCENTE CALIFICADOR

DOCENTE CALIFICADOR

## **DEDICATORIA**

*Dedicado a mis padres, mi hermano y familia en general por todo el apoyo brindado en todas las etapas de mi vida.*

*Jaime Andrés Cárdenas Córdova*

## **AGRADECIMIENTO**

*A la vida y a mi familia que son la razón  
impulsadora para seguir labrando mi futuro  
personal y profesional.*

*Jaime Andrés Cárdenas Córdova*

# ÍNDICE DE CONTENIDOS

ÍNDICE DE CONTENIDOS	VI
<b>ÍNDICE DE GRÁFICOS Y FIGURAS</b>	<b>IX</b>
<b>ÍNDICE DE TABLAS</b>	<b>XII</b>
RESUMEN EJECUTIVO	XIII
INTRODUCCIÓN	XIV
CAPÍTULO I	1
1 El Problema	1
1.1. Tema	1
1.2. Planteamiento del Problema	1
1.2.1. Contextualización	1
1.2.2. Análisis Crítico	3
1.2.3. Prognosis	4
1.2.4. Formulación del Problema	5
1.3. Preguntas Directrices	5
1.4. Delimitación del Problema	5
1.4.1. Declaración de las Variables	5
1.5. Justificación	6
1.6. Objetivos	7
1.6.1. Objetivo General	7
1.6.2. Objetivos Específicos	7
CAPÍTULO II	8
2. Marco Teórico	8
2.1. Antecedentes Investigativos	8
2.2. Fundamentación Legal	8
2.3. Categorías Fundamentales	9
2.3.1. Componentes del Sistema de Codificación de Voz a Texto	9
2.3.1.1. La Voz Humana	9
2.3.1.1.1. Fonología y Fonética	12
2.3.1.1.2. Formantes	12
2.3.1.2. Micrófono	15
2.3.1.2.1. Tipos de Micrófonos	17

2.3.1.3.	Matlab	20
2.3.1.3.1.	Aplicaciones Típicas de Matlab	21
2.3.1.4.	Digitalización de la Voz	21
2.3.1.4.1.	Muestreo y Cuantificación	22
2.3.1.4.2.	Muestreo Periódico	23
2.3.1.4.3.	Teorema de Nyquist	24
2.3.1.4.4.	Cuantificación	25
2.3.1.5.	Teoría de Fourier	27
2.3.1.5.1.	Transformada de Fourier	27
2.3.1.6.	Personas Discapacitadas	28
2.3.1.6.1.	Introducción	28
2.3.1.6.2.	Que Significa Discapacidad	29
2.3.1.7.	Deficiencias Sensoriales	31
2.3.1.7.1.	Deficiencia Auditiva	31
2.3.1.7.2.	Grado de Pérdida	32
2.3.1.7.3.	Edad de Inicio	33
2.3.1.7.4.	Causas	33
2.3.1.7.5.	Mejora de la Audición Residual Utilizando Equipo	34
2.3.1.7.6.	Desarrollo de Técnicas de Comunicación	35
2.3.1.7.7.	Adaptar Evaluación y Entrenamiento	35
2.4.	Hipótesis	35
CAPÍTULO III		36
3.	Marco Metodológico	36
3.1.	Enfoque	36
3.2.	Modalidad Básica de Investigación	36
3.2.1.	Bibliografía Documental	36
3.2.2.	De Campo	37
3.2.3.	Aplicada	37
3.3.	Tipo o Nivel de Investigación	37
3.3.1.	Exploratoria	37
3.3.2.	Descriptiva	37
3.4.	Población y Muestra	38
3.5.	Métodos y Técnicas de la Investigación	38



3.6.	Operacionalización de las Variables	38
3.7.	Plan de Recolección de Información	42
CAPÍTULO IV		44
4.	Análisis de Resultados	44
CAPÍTULO V		51
5.	Conclusiones y Recomendaciones	51
5.1.	Conclusiones	51
5.2.	Recomendaciones	51
CAPÍTULO VI		52
6.	Propuesta	52
6.1.	Datos Informativos	52
6.2.	Antecedentes de la Propuesta	52
6.3.	Justificación	53
6.4.	Objetivos	54
6.4.1.	Objetivo General	54
6.4.2.	Objetivos Específicos	54
6.5.	Análisis de Factibilidad	55
6.6.	Fundamentación	56
6.6.1.	Transformada de Fourier	57
6.6.1.1.	Transformada Discreta de Fourier	57
6.6.1.2.	Transformada Rápida de Fourier	59
6.6.2.	Filtrado de Señales	60
6.6.2.1.	Tipos de Filtros	60
6.6.3.	Evaluación del Sistema de Transformación de Señales de Voz	62
6.6.4.	Parámetros para evaluar la calidad de transformar las Señales de Voz	62
6.7.	Metodología	63
6.7.1.	Proceso de Adquisición de la Señal del Hablante	63
6.7.2.	Proceso de Filtrado de la Señal	64
6.7.2.1.	Filtro Pasa Altas	64
6.7.2.2.	Filtro de Pre-Énfasis	65
6.7.3.	Normalización	66
6.7.4.	Transformada de la señal	66

6.7.5.	Diseño del Sistema de Codificación de Señales de Voz a Texto utilizando la interfaz guide de Matlab	67
6.7.5.1.	Diseño del Sistema	67
6.7.5.2.	Normalización de Señales	70
6.7.5.3.	Cálculo de la FFT de la Señal	71
6.7.5.4.	Cálculo de Proceso de Error	72
6.7.6.	Modo de Uso	74
6.7.6.1.	Requisitos de Hardware y Software	74
6.7.6.2.	Pasos para el uso de la aplicación	75
6.7.7.	Ejemplos de Transformación de Señales de Voz a Texto	81
6.7.8.	Análisis e Interpretación de Resultados	83
6.7.8.1.	Resultados de la Evaluación en Ambientes Ideales	83
6.7.8.2.	Resultados de la Evaluación en Ambientes con Ruido Moderado	85
6.7.8.3.	Resultados de la Evaluación con Alto Ruido	86
CAPÍTULO VII		87
7.	Conclusiones y Recomendaciones	87
7.1.	Conclusiones	87
7.2.	Recomendaciones	88
Bibliografía		89
ANEXOS		91

## ÍNDICE DE GRÁFICOS Y FIGURAS

Figura 1.1 Árbol del Problema.....	3
Figura 2.1 Categorías fundamentales.....	9
Figura 2.2 Corte Esquemático del Aparato Fonatorio Humano.....	10
Figura 2.3 Cuadro de representación del Sistema vocal humano .....	11
Figura 2.4 Tablas de gamas de niveles de intensidad y de frecuencias de voz.....	11
Figura 2.5 Gráfica de formantes en la voz humana .....	13
Figura 2.6 Gráfica de la Pronunciación de las vocales .....	14
Figura 2.7 Representación frecuencial de las formantes con las vocales .....	15
Figura 2.8 Patrón Omnidireccional.....	17
Figura 2.9 Patrón Bidireccional .....	18
Figura 2.10 Patrón Cardioide .....	18
Figura 2.11 Patrón Hipercardioide.....	18
Figura 2.12 Entorno gráfico de Matlab.....	21
Figura 2.13 Proceso de Muestreo.....	23
Figura 2.14 Gráfica de Conversor Continuo/Discreto .....	24
Figura 2.15 Gráfica de Proceso de Cuantificación.....	25
Gráfico 4.1 Resultados - Pregunta 1 .....	44
Gráfico 4.2 Resultados - Pregunta 2 .....	45
Gráfico 4.3 Resultados - Pregunta 3 .....	46
Gráfico 4.4 Resultados - Pregunta 4 .....	47
Gráfico 4.5 Resultados - Pregunta 5 .....	48
Gráfico 4.6 Resultados - Pregunta 6 .....	49
Gráfico 4.7 Resultados - Pregunta 7 .....	50
Figura 6.1 Tipos de Filtros.....	61
Figura 6.2 Diagrama de bloques de un sistema general de conversión de señales de voz a texto .....	63
Figura 6.3 Respuesta de un filtro pasa altas.....	65
Figura 6.4 Interfaz gráfica del sistema de transformación.....	68
Figura 6.5 Ventana de Configuración de Sonido .....	69
Figura 6.6 Señal del fonema "a" en el dominio del tiempo .....	69
Figura 6.7 Señal Filtrada.....	70

Figura 6.8 Algoritmo del proceso de normalización.....	71
Figura 6.9 Señal en el dominio de la frecuencia (FFT).....	72
Figura 6.10 Resultados del proceso de error de la vocal "a" .....	73
Figura 6.11 Vocal reconocida por el sistema de transformación.....	73
Figura 6.12 Pantalla principal de Matlab .....	75
Figura 6.13 Barra de Herramientas de Matlab .....	75
Figura 6.14 Barra de Herramientas de la sub-ventana de Matlab.....	76
Figura 6.15 Ventana de comandos de Matlab.....	76
Figura 6.16 Pantalla principal del sistema de transformación .....	77
Figura 6.17 Sección de control del sistema de transformación.....	77
Figura 6.18 Señal introducida graficada .....	77
Figura 6.19 Botón reproducir .....	78
Figura 6.20 Botón Transformar .....	78
Figura 6.21 Transformada de la señal de base de datos.....	79
Figura 6.22 Transformada de la señal introducida.....	79
Figura 6.23 Letra reconocida .....	79
Figura 6.24 Pantalla principal después del proceso de transformación .....	80
Figura 6.25 Botón salir .....	80
Figura 6.26 Ejemplo de transformación a una palabra .....	81
Figura 6.27. Ejemplo de transformación a un número .....	82
Figura 6.28. Ejemplo de transformación a una vocal .....	82

## ÍNDICE DE TABLAS

Tabla 2.1 Formantes vocálicos.....	13
Tabla 3.1 Tabla deoperacionalización de la variable independiente.....	39
Tabla 3.2 Tabla de operacionalización de la variable dependiente.....	40
Tabla 3.3 Plan de recolección de Información.....	42
Tabla 4.1 Resultados - Pregunta 1 .....	44
Tabla 4.2 Resultados - Pregunta 2 .....	45
Tabla 4.3 Resultados - Pregunta 3 .....	46
Tabla 4.4 Resultados - Pregunta 4 .....	47
Tabla 4.5 Resultados - Pregunta 5 .....	48
Tabla 4.6 Resultados - Pregunta 6 .....	49
Tabla 4.7 Resultados - Pregunta 7 .....	50
Tabla 6.1 Parámetros de la tarjeta de sonido .....	64
Tabla 6.2 Tabla de resultados en ambientes ideales .....	84
Tabla 6.3 Tabla de promedio de resultados en ambientes ideales.....	84
Tabla 6.4 Tabla de resultados en ambientes con ruido moderado.....	85
Tabla 6.5 Tabla de promedio de resultados en ambientes con ruido moderado...	86

## RESUMEN EJECUTIVO

El presente proyecto descrito a continuación, trata sobre el estudio para transformar señales de voz a texto y su aplicación como herramienta de enseñanza – aprendizaje en el Instituto de Educación Especial Dr. Camilo Gallegos, lugar donde los estudiantes con deficiencia auditiva puedan desarrollar la habilidad de reconocer los símbolos alfanuméricos utilizados en el lenguaje hablado.

Éste trabajo, además abarca los procesos más comunes pero eficientes del procesamiento de señales analógicas y ante todo de señales de voz. La Transformada Rápida de Fourier (FFT), es el procesamiento matemático esencial en el desarrollo del proyecto, permite analizar los anchos de banda de los fonemas pronunciados, para facilitar el método de transformación.

Para aplicar todas las técnicas del Procesamiento Digital de Señales, se utiliza una herramienta especializada en soluciones matemáticas, Matlab, junto con su instrumento Guide permitiendo elaborar una aplicación gráfica, didáctica y amigable con el usuario.

El desarrollo de una guía de uso permite facilitar el manejo del proyecto, bajo ciertos parámetros y normas de uso, que son descritos en respuesta a los datos arrojados durante las pruebas realizadas en el programa.

Finalmente se dan recomendaciones para tratar de incentivar a los alumnos discapacitados en el aprendizaje y reconocimiento de los símbolos utilizados en el alfabeto y números, pudiendo generar oportunidad de inclusión social y desarrollo cognoscitivo.

## INTRODUCCIÓN

En el proceso de transformación de señales de voz a texto, se debe tener en cuenta aspectos como vocalización, tono de voz, y los ambientes donde se desea realizar la transformación, los cuales deberían permanecer libres de ruidos exagerados por el motivo de afectación al programa con señales no deseadas, y esto causará la confusión de los fonemas expresados ocasionando así una mala transformación de la señal.

Los alumnos con deficiencia auditiva tendrán la posibilidad de familiarizar su lenguaje de señas, con los símbolos que describen el alfabeto, así como el de los números. Se pretende que reconozcan dichos símbolos para establecer un método de comunicación entre personas normales y personas sordo-mudas.

El propósito de la presente tesis es diseñar e implementar un sistema de codificación de señales de voz a texto para personas con deficiencia auditiva del Instituto de Educación Especial Dr. Camilo Gallegos, que servirá como material didáctico en el proceso de enseñanza – aprendizaje, para crear un nuevo método de comunicación.

El presente documento está dividido en siete capítulos, detallados de la siguiente manera:

**Capítulo I:** Descripción del problema de investigación, el mismo indica las causas y consecuencias, que justificación, y plantea los objetivos a alcanzar en el proyecto.

**Capítulo II:** Corresponde al marco teórico, se explica la parte teórica sobre procesamiento digital de señales, la técnica más común, como la Teoría de Fourier. Conceptos sobre la voz, su fuente, estructura y características, así como

los elementos para adquirir la señal de voz. Además una explicación de las deficiencias sensoriales y sus características. En este capítulo se plantea la hipótesis a ser comprobada durante el desarrollo del presente proyecto y las variables de la misma.

**Capítulo III:** Se detalla el marco metodológico, explicando el enfoque, la modalidad y tipo de investigación; así como, la recolección y análisis de datos.

**Capítulo IV:** Presentación y análisis de los resultados de la encuesta realizada a los docentes del Instituto de Educación Especial Dr. Camilo Gallegos con el fin de evaluar la factibilidad del desarrollo de un sistema de codificación de señales de voz a texto

**Capítulo V:** Se presentan las conclusiones y las recomendaciones luego del análisis de los resultados obtenidos de la encuesta.

**Capítulo VI:** Trata sobre la descripción del desarrollo de la propuesta, describe los procesos de codificación y análisis de señales mediante la aplicación de filtrado de señales, normalización de datos y transformada rápida de Fourier (FFT), así como la implementación en la herramienta gráfica Guide de Matlab. Se analiza e interpreta los datos obtenidos en cada proceso del sistema de transformación de señales de voz a texto. Se detalla una guía de uso del sistema para facilitar la manipulación del mismo.

**Capítulo VII:** Presentación de las conclusiones y recomendaciones del proyecto una vez realizado el análisis de los resultados obtenidos en el capítulo anterior.

En la parte final del proyecto se presentan la bibliografía y los anexos.



## **CAPÍTULO I**

### **1. EL PROBLEMA:**

#### **1.1. Tema:**

“Sistema de codificación de señales de voz a texto usando Matlab para personas con deficiencia auditiva del Instituto de Educación Especial Dr. Camilo Gallegos”

#### **1.2. Planteamiento del Problema**

##### **1.2.1. Contextualización:**

El ser humano a través de la historia ha sentido la necesidad de comunicarse, por tal necesidad ha creado muchos sistemas de comunicación, empezando por el telégrafo, teléfono, radio, televisión, y por último el celular, pero todos estos inventos han sido únicamente desarrollados para personas normales sin discapacidades.

La comunicación se puede dar cuando exista un transmisor o emisor, un medio por el cual se envíe la información y un receptor, pero si uno de estos elementos falla o no funciona, la comunicación no existe. Este problema se ha dado a lo largo de los años con personas que tienen deficiencia auditiva, evidenciando su exclusión de la sociedad.

Los métodos que se han desarrollado para posibilitar a las personas con deficiencia auditiva, son netamente físicos, como por ejemplo la lectura de labios

y el lenguaje de señas, pero dichos métodos implican otros problemas en la comunicación. Para el primer método, la persona discapacitada debe saber de antemano el idioma que habla la otra persona, caso contrario la lectura de labios va a ser mal interpretada. En el segundo método la persona no discapacitada deberá conocer el lenguaje de señas y dominarlo para poder desarrollar una comunicación adecuada.

Los métodos descritos anteriormente deben realizarse siempre en presencia de ambas personas y a corta distancia, viéndose privadas de hacer uso de la tecnología donde se emitan señales de audio como: radio, teléfonos celulares y fijos, etc. Por lo tanto se han visto excluidos y marginados de muchas actividades sociales.

En la actualidad, en nuestro país, se están implementando planes de inclusión social, en donde se pretende ayudar a todas las personas que tengan cualquier tipo de discapacidad física o mental, obligando a las empresas a contratar personas con discapacidades, y sean ubicadas en áreas en donde son aptas, y puedan desarrollar su trabajo con total normalidad. De alguna manera se está tratando de eliminar la marginación y discriminación que ha existido contra las personas discapacitadas.

A nivel nacional existen 35274 personas con deficiencia auditiva, según los datos del CONADIS. Alrededor de 1488 personas forman parte de las estadísticas y base de datos del Consejo Nacional de Discapacitados de la Provincia de Tungurahua, solo dichas personas pueden tener acceso a algunos beneficios.

En la ciudad de Ambato existe un centro de educación especial para personas con discapacidad auditiva, en donde se imparten asignaturas de educación básica mediante el lenguaje de señas, sin embargo la mayor parte de personas sordomudas son analfabetas por falta de recursos o de información.

En la actualidad se están llevando a cabo programas de inclusión social para ayudar a personas discapacitadas, pero mientras no existan medios o formas de comunicación didáctica, diseñada para dichas personas, no podrán tener una verdadera inclusión, y seguirán siendo víctimas de la discriminación.

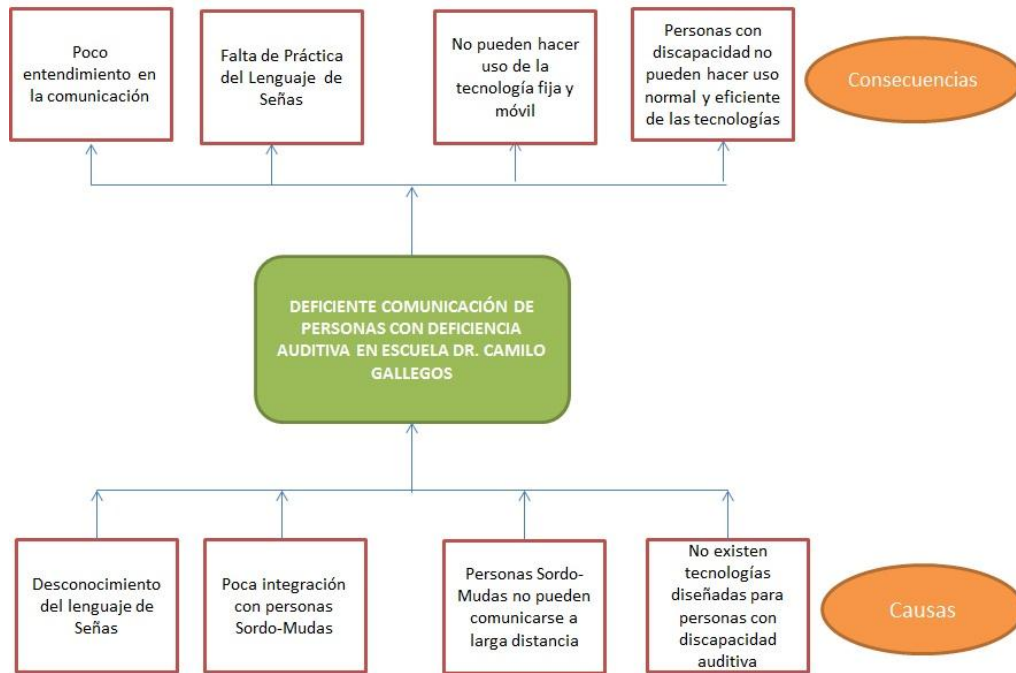


Figura 1.1 Árbol del problema  
Elaborado por: Investigador

### 1.2.2. Análisis crítico

La vida de un ser humano se da en base a la comunicación, el interactuar con otras personas es parte esencial para su desarrollo, pero esto no sucede cuando uno de los entes de la comunicación no tiene la capacidad de establecer un tipo de comunicación.

A partir de éste problema se han tratado de dar soluciones, las cuales son muy útiles cuando se trata de la comunicación entre personas con discapacidad auditiva, pero para una persona normal que no tenga, por ejemplo, conocimiento del lenguaje de señas, va a existir poco entendimiento en el proceso de comunicación entre dichas personas.

Por tal razón siempre ha existido un tipo de discriminación hacia las personas con discapacidad auditiva; lo cual acentúa en la falta de interés por parte de las personas normales, en el aprendizaje de los lenguajes de comunicación de las personas sordo-mudas.

Esto acarrea otros inconvenientes, una persona con discapacidad auditiva no puede establecer una comunicación a larga distancia con otras personas, debido que su comunicación es corta (en presencia de otra persona), por tal razón se les hace imposible hacer uso de la telefonía fija y móvil que en la actualidad son necesidades básicas. Todas las nuevas tecnologías apuntan a formas de comunicación, pero todas desarrolladas para personas normales, no aptas para personas con discapacidad auditiva. Dichas personas no pueden integrarse con normalidad y de manera eficiente, por no existir medios diseñados exclusivamente para personas discapacitadas donde tengan la facilidad de desenvolverse con naturalidad.

### **1.2.3. Prognosis**

De continuar con dichos problemas, las personas con deficiencia auditiva, no podrán establecer ningún tipo de comunicación con el resto de personas que no manejen sus métodos de comunicación, por tanto, no les será fácil adentrarse a la sociedad con normalidad. Además no podrán conseguir empleos en donde se puedan desenvolver con facilidad y normalidad, en consecuencia seguirán siendo marginados por la sociedad.

Por otro lado las personas se verían privadas o por lo menos limitadas en su acceso al conocimiento, lo cual cerraría un círculo de creciente marginación social y económica.

#### **1.2.4. Formulación del problema**

¿Cómo facilitar la comunicación de personas con deficiencias auditivas usando un sistema de codificación de señales de voz a texto?

#### **1.3. Preguntas directrices**

- ¿Habrá la posibilidad de analizar las características de las personas con deficiencia auditiva?
- ¿Es posible desarrollar el sistema de codificación de señales de voz a texto en el programa Matlab?
- ¿Se podrá plantear una propuesta que permita mejorar la comunicación de las personas con deficiencia auditiva mediante la implementación de un sistema de codificación de señales de voz a texto usando Matlab?

#### **1.4. Delimitación del Problema**

##### **Contenido:**

- **Área:** Procesamiento Digital de Señales
- **Aspecto:** Codificación de señales de voz a texto
- **Campo:** Electrónica

##### **1.4.1. Declaración de las Variables**

**Variable Independiente:** Sistema de Codificación de señales de voz a texto usando Matlab.

**Variable Dependiente:** Capacidad de comunicación de Personas con Deficiencia Auditiva.

## **1.5. Justificación**

Este proyecto es importante porque se está utilizando tecnología que ha existido hace algunos años, pero que ha sido utilizada en otros tipos de aplicaciones como digitalización de señales de audio, encriptación de información, telefonía fija, telefonía celular, etc.

La aplicación que se le va a dar al procesamiento digital de señales de voz, está dada en función de las necesidades de las personas con discapacidad auditiva, que a lo largo de la historia se les ha tratado de buscar métodos y formas, por las cuales, dichas personas puedan establecer algún tipo de comunicación con el resto de la sociedad.

El problema de comunicación de las personas con deficiencia auditiva es de mucha importancia al momento de la integración e interacción social. La creación de un nuevo método comunicacional les permitirá desarrollarse de una manera más normal dentro de la sociedad, permitiendo a una persona no discapacitada aprender el método clásico de comunicación de personas sordo-mudas, como es el lenguaje de señas.

Una aplicación didáctica beneficiará en gran medida a todo el universo de personas que tengan discapacidad auditiva, con su uso impartido dentro de la institución educativa tendrán la posibilidad de combinar su método de comunicación originario con el nuevo dentro de los procesos de aprendizaje.

En consecuencia se logrará un mayor impulso tecnológico, debido que se trataría de desarrollar una nueva aplicación con técnicas modernas, y de fácil uso. Tratando de que no sea necesario tener un conocimiento previo de tecnología moderna para hacer uso del nuevo sistema de comunicación.

## **1.6. Objetivos**

### **1.6.1. Objetivo general**

Sistema de codificación de señales de voz a texto, usando Matlab para personas con deficiencia auditiva del Instituto de Educación Especial Dr. Camilo Gallegos.

### **1.6.2. Objetivos Específicos**

- Analizar el proceso de codificación de señales de voz a texto
- Establecer las características de las deficiencias auditivas de las personas con sordera en el Instituto de Educación Especial Dr. Camilo Gallegos.
- Plantear una propuesta que permita mejorar la comunicación de las personas con deficiencia auditiva mediante la implementación de un sistema de codificación de señales de voz a texto usando matlab.

## **CAPÍTULO II**

### **2. MARCO TEÓRICO:**

#### **2.1. Antecedentes Investigativos:**

Se ha consultado en la biblioteca de la Facultad de Ingeniería en Sistemas, Electrónica e Industrial temas relacionados con la “Codificación de Señales de Voz a texto”, pero no se encontró ninguna tesis o proyecto con dicho tema o similar.

En otras universidades se encuentran tesis con aplicaciones de codificación de señales de voz pero dirigidas a prácticas de domótica y robots controlados por voz como es el caso de la tesis: “SISTEMA DE CONTROL ELECTRÓNICO POR COMANDOS DE VOZ PARA PROTOTIPO DE ROBOT COMPAÑERO”, desarrollado por el Ing. Jendry Manuel Villacís de la Universidad Politécnica de Chimborazo, que servirá como guía y material de consulta para el desarrollo de la presente tesis.

#### **2.2. Fundamentación Legal:**

El proyecto está basado y sustentado bajo la Constitución del Ecuador, en la sección de Personas Discapacitadas, las cuales están detalladas en los artículos 47, 48, y 49 en donde el estado garantiza oportunidades para las personas con discapacidad y su integración social.



## 2.3. Categorías Fundamentales:

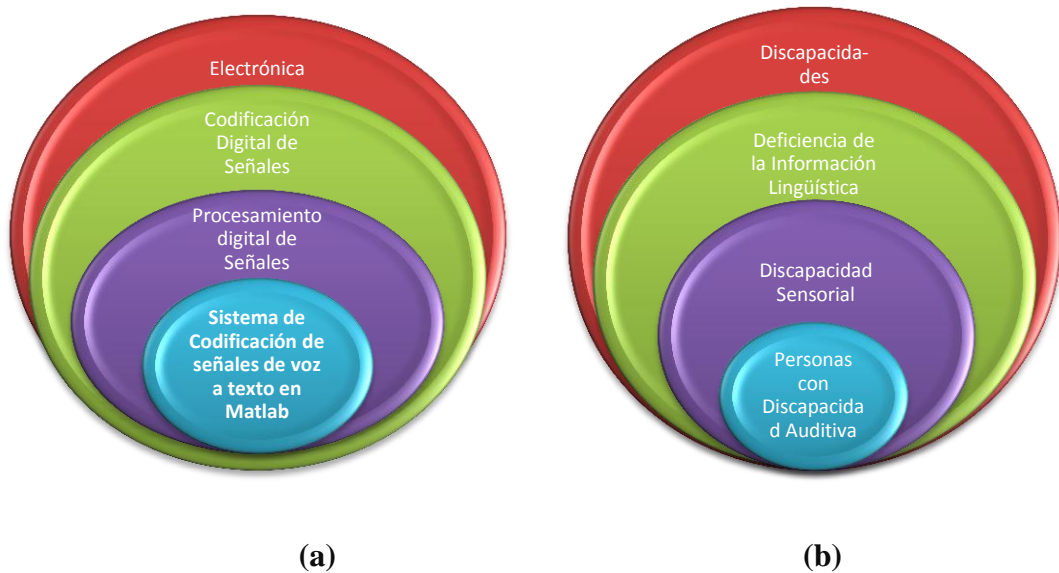


Figura 2.1 Categorías Fundamentales (a) variable independiente (b) variable dependiente

Elaborado por: Investigador

- a) **Variable Independiente:** Sistema de Codificación de señales de voz a texto usando Matlab.
- b) **Variable Dependiente:** Capacidad de comunicación de Personas con Deficiencia Auditiva.

### 2.3.1. Componentes del Sistema de Codificación de Voz a Texto

#### 2.3.1.1. La Voz Humana

“La voz humana se produce voluntariamente por medio del aparato fonatorio. Éste está formado por los pulmones como fuente de energía en la forma de un flujo de aire, la *laringe*, que contiene las *cuerdas vocales*, la *faringe*, las *cavidades*

*oral* (o bucal) y *nasal* y una serie de elementos articulatorios: los *labios*, los *dientes*, el *alvéolo*, el *paladar*, el *velo del paladar* y la *lengua*.”<sup>1</sup>

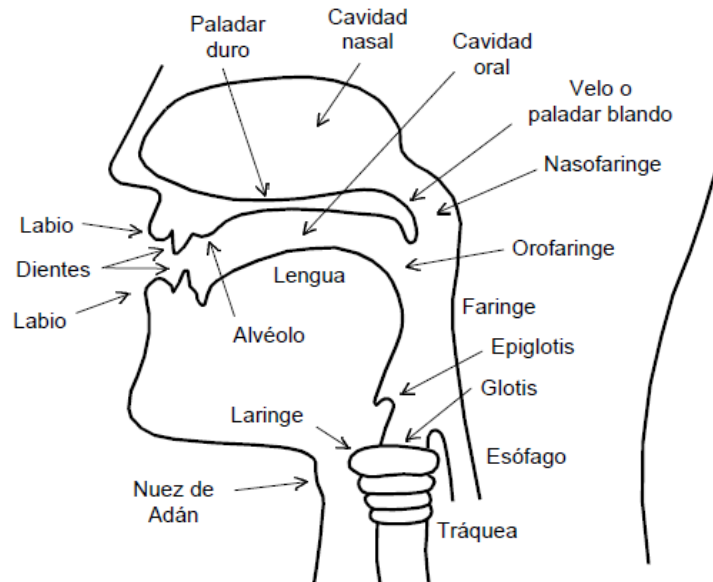


Figura 2.2. Corte Esquemático del aparato Fonatorio Humano

Fuente: [www.scribd.com](http://www.scribd.com)

La producción de sonido viene generada fundamentalmente por las cuerdas bucales, que se encuentran en la laringe. La abertura entre las cuerdas se denomina *glotis*.

A medida que se cierran las cuerdas vocales, éstas comienzan a vibrar a modo de lengüetas, produciéndose un sonido tonal, es decir periódico. La frecuencia de dicho sonido depende de varios factores, entre otros del tamaño y la masa de las cuerdas vocales, de la tensión que se les aplique y de la velocidad del flujo del aire proveniente de los pulmones. A mayor tamaño, menor frecuencia de vibración, lo cual explica por qué en los varones, cuya glotis, es en promedio mayor que la de las mujeres, por lo tanto la voz en general es más grave. A mayor tensión la frecuencia aumenta, siendo los sonidos más agudos. Así, para lograr emitir sonidos en el registro extremo de la voz es necesario un mayor esfuerzo vocal. También aumenta la frecuencia (a igualdad de las otras condiciones) al crecer la

<sup>1</sup> Miyera Federico, “LA VOZ HUMANA”

velocidad del flujo de aire, razón por la cual al aumentar la intensidad de emisión se tiende a elevar espontáneamente el tono de voz.

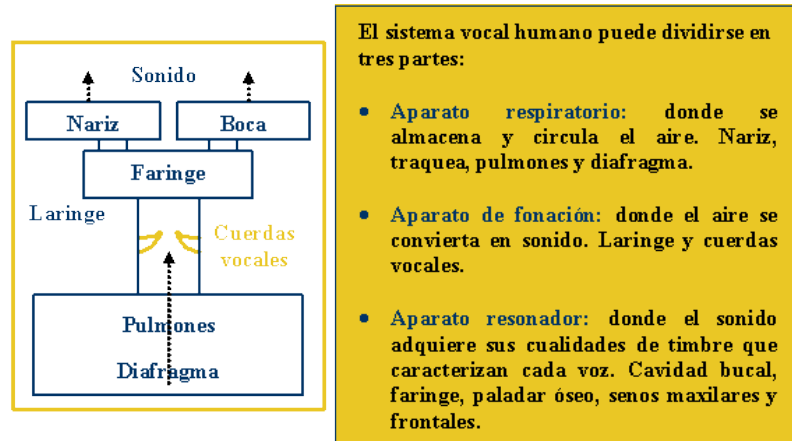


Figura 2.3 Cuadro de representación del Sistema Vocal Humano.

Fuente: www.ehu.es

A continuación en la figura, se muestra un cuadro de la gama de niveles de intensidad y sus frecuencias.

Emisión	Intensidad (w/m <sup>2</sup> )	Nivel sonoro (dB)
Nivel mínimo de la voz humana	10 <sup>-10</sup>	20
Mujer conversando en voz baja	3.16x10 <sup>-10</sup>	25
Hombre conversando en voz baja	10 <sup>-9</sup>	30
Mujer conversando en voz normal	10 <sup>-7</sup>	50
Hombre conversando en voz normal	3.16x10 <sup>-7</sup>	55
Mujer hablando en público	10 <sup>-6</sup>	60
Hombre hablando en público	3.16x10 <sup>-6</sup>	65
Mujer hablando esforzándose	10 <sup>-5</sup>	70
Hombre hablando esforzándose	3.16x10 <sup>-5</sup>	75
Mujer cantando	10 <sup>-4</sup>	80
Hombre cantando	3.16x10 <sup>-4</sup>	85
Nivel máximo de la voz humana	10 <sup>-3</sup>	90

Voz	Extensión (Hz)	Tesitura
Soprano	247-1056	SI <sub>3</sub> -DO <sub>6</sub>
Mezzosoprano	220-900	LA <sub>3</sub> -SI <sub>5</sub>
Contralto	176-840	FA <sub>3</sub> -LA <sub>5</sub>
Tenor	132-528	DO <sub>3</sub> -DO <sub>5</sub>
Baritono	110-440	LA <sub>2</sub> -LA <sub>4</sub>
Bajo	82-396	MI <sub>2</sub> -SOL <sub>4</sub>

Figura 2.4 Tablas de Gammas de Niveles de intensidad y de Frecuencias Fundamentales

Fuente: Ruiz Mezcua Belén, (2005). *La Voz y su Espectro*

### **2.3.1.1.1. Fonología y fonética**

“La Fonología se dedica al estudio de los sonidos de la lengua en cuanto a su carácter simbólico o de representación mental. A cada uno de los sonidos abstractos identificados se lo conoce como *fonema*. Uno de los objetivos de la fonología es acotar al máximo la cantidad de fonemas requeridos para representar cada idioma de una manera suficientemente precisa”.<sup>2</sup>

En tanto, la *Fonética*, se refiere a los sonidos en el habla, incluyendo su producción acústica y los procesos físicos y fisiológicos de emisión y articulación involucrados. Se ocupa, por consiguiente, de las realizaciones de los fonemas.

### **2.3.1.1.2. Formantes**

“Los formantes son elementos que sirven para distinguir componentes del habla humana, principalmente, las vocales y sonidos sonantes. El formante con la frecuencia más baja se llama F1, el segundo F2, el tercero F3, etc.”.<sup>3</sup>

Habitualmente, son necesarios sólo los dos primeros para caracterizar una vocal, aunque la pueden caracterizar más formantes. Generalmente, los formantes posteriores determinan propiedades acústicas como el timbre.

Los dos primeros formantes se determinan principalmente por la posición de la lengua. Sucediendo que F1 tiene una frecuencia más alta cuanto mas baja esta la lengua, es decir una mayor abertura. Para el F2 tiene mayor frecuencia cuanto mas hacia delante esta posicionada la lengua.

---

<sup>2</sup> Miyera Federico, LA VOZ HUMANA.

<sup>3</sup> Velásquez Ramírez Genoveva, *Sistema de Reconocimiento de Voz*, (2008).

No todos los sonidos se componen de formantes definidos. Solamente aparecen en sonantes, que incluyen los sonidos pulmonares: vocales, aproximantes y nasales. Éstas últimas tienen un formante adicional F3, en torno a los 1500 Hz.

Si la frecuencia fundamental es mayor que la frecuencia de los formantes, entonces el carácter del sonido se pierde y se vuelven difíciles de distinguir, por lo cual son difíciles de reconocer.

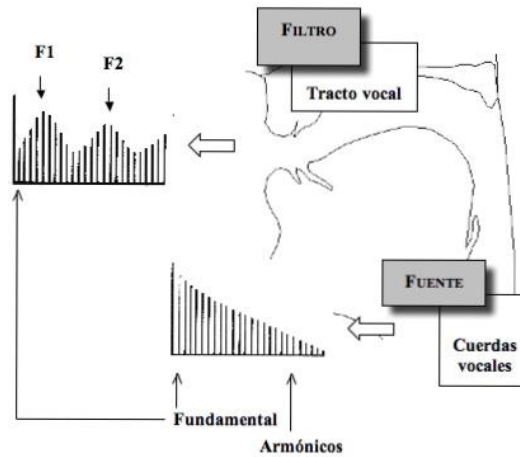


Figura 2.5 Gráfica de formantes en la voz humana

Fuente: <http://dravalles.blogspot.com/>

A continuación se muestran algunos anchos de banda entre los cuales se localizan las vocales:

Tabla 2.1 Formantes Vocálicos

Fuente: <http://es.wikipedia.org/wiki/Formante>

Formantes Vocálicos	
Vocal	Región principal formantica
/u/	200 a 400 Hz
/o/	400 a 600 Hz
/a/	800 a 1200 Hz
/e/	400 a 600 y 2200 a 2600 Hz
/i/	200 a 400 y 3000 a 3500 Hz

## Clasificación de las Vocales <sup>4</sup>

Por el modo de articulación (formante 1)

- Cerradas (i, u)
- Medias (e, o)
- Abiertas (a)

Por el lugar de articulación (formante 2)

- Anteriores (i, e)
- Centrales (a)
- Posteriores (o, u)

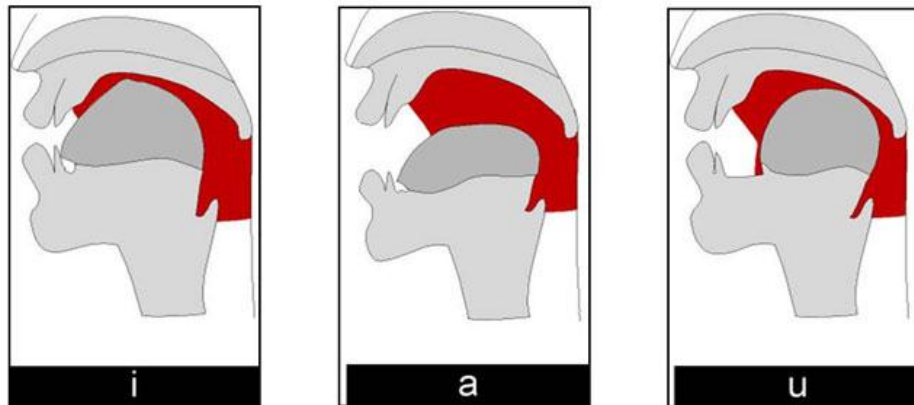


Figura 2.6 Gráfica de la Pronunciación de las Vocales

Fuente: <http://cantoyvocalizacion.blogspot.com/>

### Triángulo Vocálico

A continuación se muestra una figura en donde se explica en que frecuencias se generan las formantes al pronunciar las vocales.

---

<sup>4</sup> Fuente: <http://agamenon.tsc.uah.es/ModeloVoz>

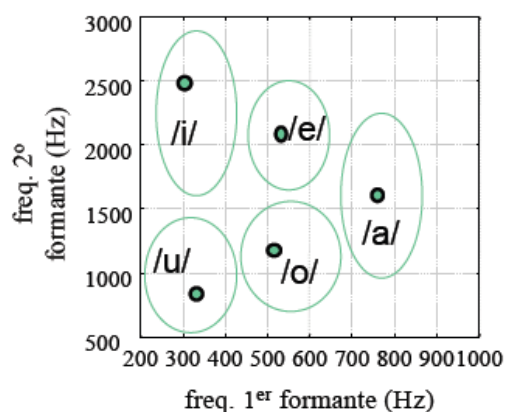


Figura 2.7 Representación frecuencial de las formantes con las vocales

Fuente: <http://agamenon.tsc.uah.es/ModeloVoz>

### 2.3.1.2. Micrófono

Es todo aquel elemento que transforma energía acústica (sonido) en energía eléctrica (señal de audio). Es una variante de una serie de transductores. La fidelidad con la que este elemento genera una representación eléctrica del sonido depende en parte del método que utiliza para la conversión de energía. A lo largo del tiempo se ha diseñado una serie de métodos distintos para propósitos varios; de esta manera, hoy podemos contar con una amplia variedad en tipos de micrófono disponibles.

La calidad de cada micrófono viene dada por sus características, las cuales se describen a continuación:

**1. Sensibilidad:** es la eficiencia del micrófono, la relación entre la presión sonora que incide (expresada en Pascales) y la tensión eléctrica de salida (expresada en voltios). O sea, expresa que tan bien convierte el micrófono la presión acústica en voltaje de salida. La sensibilidad se expresa en mili-voltios por Pascal.

Al utilizar el mili-voltio, la sensibilidad puede ser representada en un voltímetro de la siguiente manera: a mayor voltaje, mayor sensibilidad.

**2. Fidelidad:** indica la variación de sensibilidad con respecto a la frecuencia. Además, la fidelidad, viene definida como la respuesta en frecuencia del micrófono, cuanto mas lineal sea la respuesta en frecuencia mayor fidelidad tendrá el micrófono. La fidelidad se expresa en dB.

En función de esta respuesta en frecuencia o fidelidad se elabora la llamada Curva de respuesta de un micrófono, que es la representación gráfica del nivel obtenido en la captación de sonidos de igual intensidad, pero de distinta frecuencia.

**3. Directividad:** esta característica determina en que dirección capta mejor el sonido un micrófono, es decir indica la sensibilidad del micrófono a las diferentes direcciones.

**4. Ruido de fondo:** es la tensión que entrega el micrófono sin que exista ningún sonido incidiendo sobre él. Este ruido se produce por el movimiento térmico de los electrones en la carcasa que no tiene masa. El ruido de fondo debe estar en torno a los 60dB, pero mientras más bajo sea, mejor calidad ofrece el micrófono.

**5. Impedancia interna:** es la resistencia que opone el micrófono al paso de la corriente. La impedancia según su valor viene caracterizada por baja, alta y muy alta impedancia.

- Lo-Z Baja impedancia (alrededor de 200 Ohmios)
- Hi-Z Alta impedancia (1 K Ohmios o 3 K Ohmios e incluso 600 Ohmios)
- VHi-Z Muy alta impedancia (mas de 3 K Ohmios).

Si el micrófono es de alta impedancia y se tiene un cable largo se produce una pérdida muy grande. Si se tiene una impedancia baja se puede utilizar



un cable muy largo y no se pierde tanto la señal. Por último, se puede bajar la resistencia para evitar pérdidas en altas frecuencias.<sup>5</sup>

### 2.3.1.2.1. Tipos de Micrófonos<sup>6</sup>

Se pueden dividir según:

#### a) La Directividad

Como ya se mencionó en las características, según la directividad hay tres tipos de micrófonos

:

1. Micrófono omnidireccional: este es aquel cuyo diagrama polar es considerado como un círculo perfecto. Aunque esto es en el caso ideal.

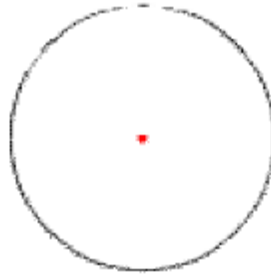


Figura 2.8 Patrón Omnidireccional

Fuente: [www.sonidoyaudio.com](http://www.sonidoyaudio.com)

2. Micrófono bidireccional: en este caso el diagrama polar muestra como captan por la parte frontal y la lateral.

---

<sup>5</sup> *Micrófonos*, Curso Intensivo de sonido EMSIA, (2002).

<sup>6</sup> [www.sonidoyaudio.com](http://www.sonidoyaudio.com)

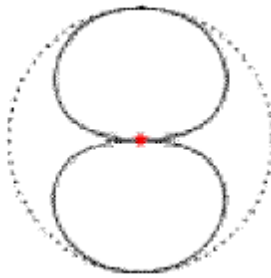


Figura 2.9 Patrón Bidireccional

Fuente: Fuente: [www.sonidoyaudio.com](http://www.sonidoyaudio.com)

3. Micrófono unidireccional: existe gran número de patrones polares para este tipo.

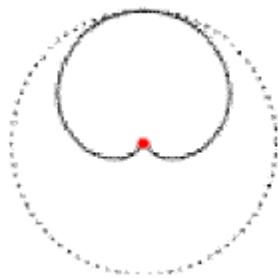


Figura 2.10 Patrón Cardioide

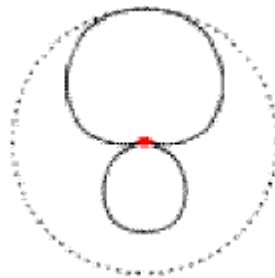


Figura 2.11 Patrón Hipercardioide

Fuente: Fuente: [www.sonidoyaudio.com](http://www.sonidoyaudio.com)

## b) El Transductor

Se encuentran tres grandes grupos según el tipo de transductor.

- **Micrófono de Condensador o Capacitor:** lo que sucede con este micrófono es que las ondas sonoras provocan el movimiento oscilatorio del diafragma, el cual actúa como una de las placas de un capacitor y la vibración provoca una variación en la energía almacenada en el condensador que forma el núcleo de la capsula microfónica. Esta variación genera una tensión eléctrica que es la señal

de salida del sistema. La señal de salida de este sistema es análoga. Según el principio de operación:

- ✓ Micrófono de condensador DC.
- ✓ Micrófono de condensador electret.
- ✓ Micrófono de condensador de radiofrecuencia (RF).

Para nuestra aplicación en MATLAB utilizamos un micrófono simple de tipo electret, ya que son los menos costos y más comunes.

- **Micrófono Dinámico:** trabajan por medio de inducción electromagnética, la vibración del diafragma provoca el movimiento de una bobina móvil o cinta corrugada ancladas a un imán permanente que genera un campo magnético que a su vez genera una tensión eléctrica, que es la señal de salida. Esta señal eléctrica es análoga. Hay dos tipos básicos:

- ✓ Micrófono de bobina móvil o dinámico
- ✓ Micrófono de cinta

- **Micrófono piezoeléctrico:** utilizan el fenómeno de piezoelectricidad, cuando las ondas sonoras hacen vibrar el diafragma el movimiento de este hace que se mueva el material contenido en su interior (cuarzo, sales de Rochéllé, carbón, etc.). La fricción entre estas partículas generan sobre la superficie del material una tensión eléctrica.

La respuesta en frecuencia de estos micrófonos es muy irregular, ya que son micrófonos piezoeléctricos:

- ✓ Micrófono de carbón
- ✓ Micrófono de cristal
- ✓ Micrófono de cerámica

### 2.3.1.3. Matlab

MATLAB es el nombre abreviado de “MAT rix LABoratory”. MATLAB es un programa para realizar cálculos numéricos con vectores y matrices. Como caso particular puede también trabajar con números escalares tanto reales como complejos, con cadenas de caracteres y con otras estructuras de información más complejas. Una de las capacidades más atractivas es la de realizar una amplia variedad de gráficos en dos y tres dimensiones, y además posee un lenguaje de programación propio.

MATLAB es un gran programa de cálculo técnico y científico. Para ciertas operaciones es muy rápido, cuando puede ejecutar sus funciones en código nativo con los tamaños más adecuados para aprovechar sus capacidades de vectorización. MATLAB dispone de un código básico y de varias librerías especializadas (toolboxes).

El lenguaje esta construido por código llamado M-code que puede ser fácilmente ejecutado en la ventana de comandos. Con lo cual se pueden crear funciones, etc. Pero la razón principal para la elección de este lenguaje de programación son las herramientas que proporciona para el procesamiento de señales, y el conjunto de funciones para el procesamiento digital.

Además, para crear entornos gráficos se puede utilizar el GUIDE de MATLAB, que provee herramientas para crear GUIs, ‘Graphical User Interface’, con lo cual se puede crear la forma del entorno grafico así como asociar funciones a los elementos del GUI. MATLAB también incluye funciones para manipular archivos.<sup>7</sup>

---

<sup>7</sup> García de Jalón, *Aprenda Matlab 7.0*, (2005).

### 2.3.1.3.1. Aplicaciones Típicas de Matlab <sup>8</sup>

- ✓ Cálculo matemático
- ✓ Desarrollo de algoritmos
- ✓ Adquisición de datos
- ✓ Modelado, simulación y prototipado
- ✓ Análisis de datos y visualización
- ✓ Gráficos
- ✓ Desarrollo de aplicaciones e interfaces gráficas de usuario (GUI)

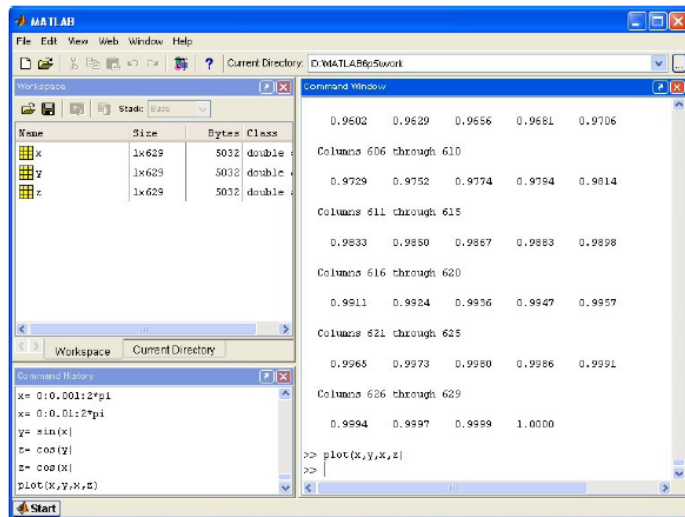


Figura 2.12 Entorno Gráfico Matlab

Fuente: Ramirez Perez Javier, *Introducción a Matlab y Simulink*.

### 2.3.1.4. Digitalización de la Voz

Las señales se representan matemáticamente como funciones de una o más variables independientes. La variable independiente más común es el tiempo, y algunas señales que dependen de él son, por ejemplo, la voz, una onda de radio, un electrocardiograma, etc.

<sup>8</sup> Ramirez Perez Javier, *Introducción a Matlab y Simulink*.

Otras señales, tales como las imágenes, son funciones de 2 variables independientes, ya que contienen información de brillo o de colorido en función de las coordenadas X e Y de un plano.

Digitalizare es convertir la entrada de voz a una forma que el reconocedor pueda procesar o que la señal sea más accesible para manipular posteriormente.

Las redes digitales de transmisión de voz y datos son comunes en nuestra era. Fueron creadas ya que presentan ciertas ventajas sobre las redes analógicas como por ejemplo: que conservan la señal casi inalterable a través de su recorrido. Es decir que es más difícil que la comunicación se vea afectada por factores externos como el ruido eléctrico.

En definitiva, digitalizar una señal de voz no es otra cosa que tomar muestras (a intervalos de tiempo regulares) de la amplitud de la señal analógica y transformar esta información a binario. Este proceso se denomina muestreo.

Una vez digitalizada la señal, se continúa con el **Procesamiento de Señales**, lo cual es un área de la Ingeniería Electrónica que se concentra en la representación, transformación, manipulación de señales, y de la información que ellas contienen.

#### **2.3.1.4.1. Muestreo y Cuantificación**

Muestreo consiste en el proceso de conversión de señales continuas a señales discretas en el tiempo, es un paso para digitalizar una señal analógica. Este proceso se realiza midiendo la señal en momentos periódicos del tiempo.

Lo más común es encontrarnos señales en tiempo discreto como consecuencia del muestreo en tiempo continuo. Bajo ciertas condiciones, una señal en tiempo continuo se puede representar de forma exacta mediante muestras tomadas en instantes discretos de tiempo.

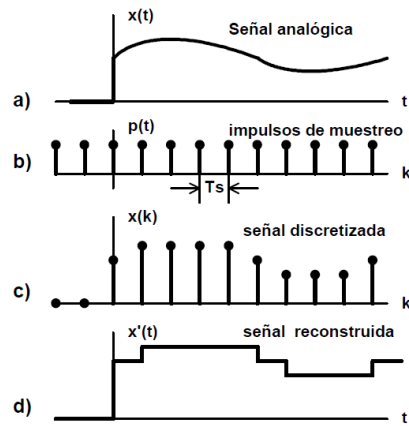


Figura 2.13 Proceso de Muestreo

Fuente: Velásquez Ramírez Genoveva, *Sistema de Reconocimiento de Voz*, (2008).

#### 2.3.1.4.2. Muestreo Periódico

La forma típica de obtener una representación discreta  $x[n]$  a partir de una señal continua  $x_c(t)$  es mediante muestreo periódico por medio de un conversor continuo/discreto (C/D), obteniendo la secuencia discreta según la expresión:

Ec. 1

$$x[n] = x_c(nT), \quad -\infty < n < \infty$$

Siendo  $T$  el periodo de muestreo, y  $f_s=1/T$  la frecuencia de muestreo.

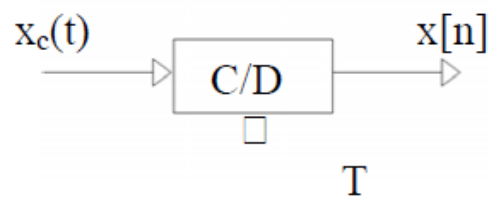


Figura 2.14 Gráfica de Conversor Continuo/Discreto

Fuente: Elaborado por el Investigador

El sistema que realiza la operación de la figura 21 se denomina conversor ideal de tiempo continuo a discreto (C/D). En la práctica esta operación de muestreo se realiza mediante un conversor analógico-digital. Estos sistemas son aproximaciones al conversor C/D ideal.

### 2.3.1.4.3. Teorema de Nyquist

Si  $x[n]$  es una secuencia de muestras obtenida a partir de una señal continua en el tiempo  $x(t)$ , por medio de la relación:

**Ec. 2**

$$x[n] = x(nT), \quad \text{para } -\infty \leq n \leq \infty$$

donde  $T$  es el período de muestreo, y su recíproco,  $f_s = \frac{1}{T}$  es la frecuencia de muestreo, en muestras por segundo. También podemos expresar la frecuencia de muestreo como  $\Omega_s = \frac{2\pi}{T}$  en radianes por segundo.

Entonces el teorema de muestreo de Nyquist está definido como: sea  $x(t)$  una señal limitada en banda por:

**Ec. 3**

$$X(j\Omega) = 0 \quad \text{para } |\Omega| \geq \Omega_N$$

Entonces  $x(t)$  está únicamente determinada por sus muestras  $x[n] = x(nT)$ ,  $n = 0, \pm 1, \pm 2, \dots$  si  $\Omega_s = \frac{2\pi}{T} \geq 2\Omega_N$ .

La frecuencia  $\Omega_N$  es comúnmente referida como la frecuencia de Nyquist, y la frecuencia  $2\Omega_N$  que tiene que ser excedida por la frecuencia de muestreo es llamada la razón de Nyquist.



#### 2.3.1.4.4. Cuantificación

“En el procesamiento digital de señales sólo dispondremos de un número finito de bits para representar cada una de las muestras. Podemos representar el proceso de cuantificación mediante”<sup>9</sup>:

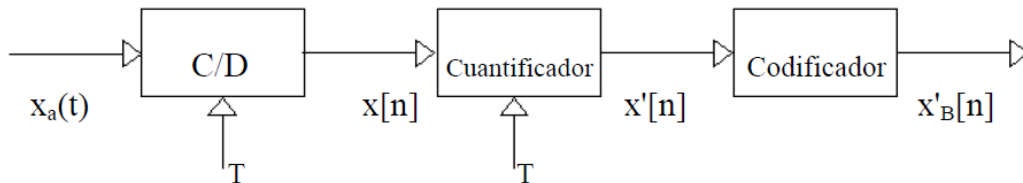


Figura 2.15 Gráfica de Proceso de Cuantificación <sup>[9]</sup>

Fuente: Borrás Gené Oriol, *Reductor de ruido mediante resta espectral en entorno Matlab*, (2006).

El cuantificador es un sistema no lineal cuyo propósito es transformar la muestra de entrada  $x[n]$  en un valor dentro de un conjunto de valores preestablecidos, representamos esta operación mediante el operador  $Q$ :

**Ec. 4**

$$x'[n] = Q\{x[n]\}$$

Se conoce como error de cuantificación (o ruido), a la diferencia entre la señal de entrada (sin cuantificar) y la señal de salida (ya cuantificada), interesa que el ruido sea lo más bajo posible. Para conseguir esto y según sea la aplicación a desarrollar, se pueden usar distintas técnicas de cuantificación <sup>[3]</sup>:

- Cuantificación uniforme
- Cuantificación logarítmica
- Cuantificación no uniforme
- Cuantificación vectorial

---

<sup>9</sup> Borrás Gené Oriol, *Reductor de ruido mediante resta espectral en entorno Matlab*. (2006)

**Cuantificación uniforme:**

En los cuantificadores uniformes o lineales la distancia entre los niveles de reconstrucción es siempre la misma, la mayoría usan un número de niveles que es una potencia de 2. No hacen ninguna suposición acerca de la señal a cuantificar, de allí que no proporcionen los mejores resultados. Pero son los mas fáciles y menos costosos a implementar.

**Cuantificación logarítmica:**

Para evitar desperdicio de niveles de reconstrucción y de ancho de banda se utiliza un método sencillo para mejorar el incremento de la distancia entre los niveles de reconstrucción conforme aumenta la amplitud de la señal. Para conseguir esto se hace pasar la señal por un compresor logarítmico antes de la cuantificación. Esta señal comprimida puede ser cuantificada uniformemente. A la salida del sistema la señal pasa por un expansor. A esta técnica se le llama compresión.

**Cuantificación no uniforme:**

Este cuantificador utiliza la función de la distribución de probabilidad, conociendo esto se puede ajustar los niveles de reconstrucción a la distribución de forma que se minimice el error cuadrático medio.

**Cuantificación vectorial:**

Este método cuantifica los datos en bloques de  $N$  muestras. En este tipo de cuantificación, el bloque de  $N$  muestras se trata como un vector  $N$ -dimensional.

### **2.3.1.5. Teoría de Fourier**

El análisis matemático de los métodos de modulación y de multicanalización de la señales utilizada en los sistemas de comunicaciones supone portadora de forma de onda senoidal y señales de información. Esto simplifica el análisis y hace predecible la operación; sin embargo, en el mundo real no todas las señales de información son senoidales. En general las señales de información son señales de voz y de vida más compleja y en esencia están compuestas de ondas senoidales de forma de muchas frecuencias y amplitudes. Las señales de información pueden tomar un número infinito de formas, incluyendo ondas rectangulares (por ejemplo, pulsos digitales), ondas triangulares, ondas de diente de sierra y otras forma no senoidales.

Estas señales requieren un enfoque no senoidal para determinar las características y el desempeño de cualquier circuito o sistema de comunicación.

Uno de los métodos más utilizados para hacerlo es el análisis de Fourier, que proporciona una forma de analizar con todo detalle el contenido de la mayoría de las señales no senoidales más compleja.

#### **2.3.1.5.1. Transformada de Fourier**

La historia del análisis de Fourier tiene más de 200 años. Sus orígenes principian unos 60 años antes del momento en que Jean Baptiste Joseph Fourier presentó la primera versión de su trabajo sobre la teoría de la conducción del calor a la Academia de París (1807).

En ese momento los esfuerzos de los físicos y matemáticos se concentraban en dos problemas principales que sentarían las bases de lo que posteriormente se conocería como análisis de Fourier: El problema de la cuerda vibrante o la propagación del sonido en un medio elástico.

La determinación de las órbitas de los planetas a partir de mediciones.

Generación de formas de onda de corriente o tensión eléctrica por medio de la superposición de senoides generados por osciladores electrónicos de amplitud variable cuyas frecuencias ya están determinadas. Análisis en el comportamiento armónico de una señal, y reforzamiento de señales.

Estudio de la respuesta en el tiempo de una variable circuital eléctrica donde la señal de entrada no es senoidal o cosenoidal, mediante el uso de transformadas de Laplace y solución permanente senoidal en el dominio de la frecuencia.

La resolución de algunas ecuaciones diferenciales en derivadas parciales admiten soluciones particulares en forma de series de Fourier fácilmente computables, y que obtener soluciones prácticas, en la teoría de la transmisión del calor, la teoría de placas, etc.

### **2.3.1.6. Personas Discapacitadas**

#### **2.3.1.6.1. Introducción**

Se estima que más de 500 millones de personas en el mundo tienen algún impedimento físico, mental o sensorial y alrededor del 80 % de estas personas viven en los países en desarrollo.

“Generalmente a las personas con discapacidad se les niega la posibilidad de educación o de desarrollo profesional, se les excluye de la vida cultural y las relaciones sociales normales, se les ingresa innecesariamente en instituciones y tienen acceso restringido a edificios públicos y transporte debido a sus limitaciones físicas. Por si fuera poco, los discapacitados se encuentran en desventaja jurídica, ya que no cuentan con un documento oficial único que

enumere sus derechos, sino que están dispersos en una serie de dictámenes judiciales, recomendaciones de la OIT e instrumentos jurídicos.”<sup>10</sup>

Pero no solamente las personas con discapacidad sufren. La sociedad pierde la oportunidad de aprovechar el enorme potencial de las personas con discapacidad.

La ONU ha luchado por mejorar vida de las personas con discapacidad. En el decenio de 1970, el concepto de los derechos humanos de las personas con discapacidad gana más aceptación internacional. En 1971 fue aprobada por la Asamblea General la Declaración de los Derechos del Retrasado Mental y en 1975, la Declaración de los Derechos de los Impedidos, la cual fija las normas para el trato igual y el acceso igual a los servicios que aceleran la integración social de estas personas.

### **2.3.1.6.2. Que significa Discapacidad**

#### **Las Naciones Unidas y las personas con discapacidad<sup>11</sup>**

La Organización Mundial de la Salud, en su Clasificación Internacional de Deficiencias, Discapacidades y Minusvalías, distingue entre deficiencia, discapacidad y minusvalía. Estos tres conceptos aparecen definidos de la siguiente manera:

- a) Por deficiencia se entiende "toda pérdida o anormalidad de una estructura o función psicológica, fisiológica o anatómica". Las deficiencias son trastornos en cualquier órgano, e incluyen defectos en extremidades, órganos u otras estructuras corporales, así como en alguna función mental, o la pérdida de alguno de estos órganos o funciones. Algunos ejemplos de deficiencias son la ceguera,

---

<sup>10</sup> [www.cinu.org.mx](http://www.cinu.org.mx)

<sup>11</sup> [www.un.org](http://www.un.org)

sordera, pérdida de visión en un ojo, parálisis o amputación de una extremidad; retraso mental, visión parcial, pérdida del habla, mutismo.

**b)** Discapacidad significa "restricción o falta (debidas a una deficiencia) de la capacidad para realizar una actividad en la forma o dentro del margen que se consideran normales para un ser humano. Engloba las limitaciones funcionales o las restricciones para realizar una actividad que resultan de una deficiencia. Las discapacidades son trastornos que se definen en función de cómo afectan a la vida de una persona. Algunos ejemplos de discapacidades son las dificultades para ver, hablar u oír normalmente, para moverse o subir las escaleras, para agarrar o alcanzar un objeto, para bañarse, comer o ir al servicio.

**c)** Una minusvalía es una "situación desventajosa para un individuo determinado, consecuencia de una deficiencia o discapacidad, que lo limita o le impide desempeñar una función considerada normal en su caso (dependiendo de la edad, del sexo y de factores sociales y culturales). El término es también una clasificación de "las circunstancias en las que es probable que se encuentren las personas discapacitadas". "Minusvalía" describe la situación social y económica de las personas deficientes o discapacitadas, desventajosa en comparación con la de otras personas. Esta situación de desventaja surge de la interacción de la persona con entornos y culturas específicos. Se consideran minusvalías tener que permanecer postrado en la cama o confinado en casa; no poder utilizar el transporte público; estar aislado socialmente.

Las minusvalías hacen referencia a las "desventajas que experimentan los individuos como resultado de deficiencias y discapacidades; por eso, las minusvalías reflejan la interacción de los individuos con su entorno y su adaptación a él."

### **2.3.1.7. Deficiencias Sensoriales**

Las personas en su mayoría que ven y oyen sin dificultad no podrían imaginar la vida sin estas habilidades. La visión y la audición, son vitales para sobrevivir, para la salud y para la calidad de vida. “Con la pérdida de uno solo de esos sentidos, la movilidad, la comunicación y el aprendizaje resultan mucho más difíciles. Cuando ambos sentidos están dañados o se presentan discapacidades adicionales, y no somos capaces de utilizar alternativas funcionales, todos los aspectos de la vida se ven afectados.”<sup>12</sup>

Las estimaciones del número de niños con discapacidades múltiples y con daños visuales y/o auditivos, son claramente poco consistentes e imprecisas. Esto se deriva de las diferencias en las definiciones utilizadas para algunas discapacidades así como de las diferencias en los procedimientos utilizados para la obtención de las muestras. Según el Estudio Anual de Niños y Jóvenes con Deficiencias Auditivas (Wolf y Harkins, 1986), el 30,2% de los alumnos con problemas de audición, informaron que sufrían problemas adicionales, de los cuales el 9,5% tenían dos o más discapacidades.

#### **2.3.1.7.1. Deficiencia Auditiva**

La pérdida de audición puede clasificarse según el grado de la misma, edad en la que se inició, causa del daño, o estructuras afectadas. Por supuesto, existen factores como son las demandas del entorno, los otros sentidos de la persona y a menudo factores adicionales de carácter severo que tienen impacto en la deficiencia auditiva, dando como resultado habilidades y necesidades específicas. Por ejemplo, la habilidad para comprender el lenguaje oral puede diferir entre dos niños que tengan idéntica pérdida auditiva, dependerá de sus cualificaciones visuales y cognitivas así como el entorno social en el que viven.

---

<sup>12</sup> Dick Sobsey y Enid G. Wolf-Schein

### 2.3.1.7.2. Grado de Pérdida

Si una persona es *sorda*, la audición está dañada hasta tal punto que él o ella no pueden comprender el lenguaje oral sólo a través del oído, tanto si utiliza como no, aparatos auditivos. Cuando una persona es *dura de oído*, la audición está dañada hasta un punto en el que resulta difícil, pero no imposible, la comprensión del lenguaje oral sólo a través del oído, tanto sí se utilizan como no aparatos auditivos.

El grado de pérdida puede clasificarse según la frecuencia del sonido (Herzios o Hz [ciclos por segundo]) y los niveles de intensidad (Decibelios o dB) que requiere el individuo para escuchar el lenguaje oral u otros sonidos

El grado de pérdida auditiva puede diferir según las frecuencias, pero normalmente se presta atención particular a las pérdidas entre los 500-200 Hz. porque esta frecuencia es crítica para discriminar el habla.

Los individuos con *pérdidas de audición media* (Nivel 1, 35-54 dB) pueden oír la mayor parte del habla así como muchos otros sonidos de su entorno, pero en general se beneficiarían de modificaciones en las aulas, incluyendo en muchos casos la amplificación.

La logopedia y entrenamiento auditivo también pueden resultar beneficiosos para estos alumnos, aunque algunos pueden beneficiarse de las alternativas al lenguaje no oral, lo cual dependerá en parte de la naturaleza y alcance de otras discapacidades parciales. El lenguaje oral es una alternativa de recepción viable y de producción de lenguaje para muchos.

Los alumnos con *pérdidas de audición moderada* (Nivel II, 55-69 dB) requieren un lenguaje oral especial, asistencia en la audición y en el lenguaje, y amplificación, a menos que la naturaleza de la pérdida de audición impida la obtención de beneficios (la amplificación no constituye una ayuda en algunos tipos de pérdida de audición neural). Para estos alumnos, el lenguaje oral puede seguir siendo una parte importante.



Las personas con *pérdidas de audición severa* (Nivel III, 70-89 dB) y con otras discapacidades, requerirán un lenguaje oral especial, audición, lenguaje, y asistencia educativa. Algunos podrán beneficiarse de la amplificación, aunque muchos no.

Las personas con *pérdidas de audición profundas* (Nivel IV, 90 dB ó más) requieren asistencia intensiva en el lenguaje y educación. Con este nivel de daño, la alternativa de lenguaje no oral es casi siempre necesaria, sin embargo una evaluación cuidadosa llevada a cabo por profesionales de la audición y visión resulta importante para determinar el grado de intervención más apropiado.

#### **2.3.1.7.3. Edad de Inicio**

Se considera que un individuo tiene sordera pre lingüística, cuando su sordera ya estaba presente en el nacimiento o tuvo lugar antes del desarrollo del habla y lenguaje.

Las personas con sordera post-linguística son aquellos cuya sordera tuvo lugar después de haber adquirido el habla y el lenguaje. Los alumnos con discapacidades múltiples que han sido expuestos al habla y han aprendido algunas técnicas de habla receptiva y productiva básicas, tendrán mas oportunidad para adquirir un habla útil, sea cual sea el grado de pérdida, que cualquier otro alumno cuya sordera haya sido pre-lingüística

#### **2.3.1.7.4. Causas**

La sordera puede producirse por distintas causas. Por ejemplo, puede ser debida a factores ambientales como infecciones o daños antes, durante o después del nacimiento, o por la exposición a ruidos muy grandes durante un periodo de tiempo. También puede ser de origen genético, que los padres y hermanos también sean sordos.

Los niños sordos que crecen en hogares donde se utiliza la lengua de signos, a menudo aprenden una alternativa de lenguaje con más facilidad, luego éste puede

ser un factor para seleccionar una alternativa de lenguaje. La pérdida auditiva total puede estar presente desde el nacimiento o puede ser progresivo.

La frecuencia con la que tiene lugar la sordera también puede estar influida por enfermedades, y por la calidad en el cuidado de la salud. Un ejemplo dramático fue la epidemia de la rubéola 1963-1965, que triplicó el número de niños que nacieron sordos, durante ese periodo, en Estados Unidos. Desde entonces se han desencadenado diversas epidemias en Norteamérica dando como resultado un aumento en el número de individuos con pérdidas auditivas en años específicos.

#### **2.3.1.7.5. Mejora de la Audición Residual Utilizando Equipo**

El uso de audífonos adecuados deberá ser considerado siempre que se desarrollen programas de habilitación extensos. Tanto los padres como los profesores necesitan comprender, que la utilización de amplificadores puede acelerar el desarrollo de técnicas de comunicación expresivas, pero también minimizar dificultades académicas. A pesar de la amplia aceptación de este concepto, existen algunos problemas. En primer lugar, muchos niños que pueden beneficiarse de este tipo de amplificación, en especial aquellos con pérdidas auditivas medias y profundas, no se les ayuda o no lo suficientemente pronto para prevenir retrasos en el desarrollo del lenguaje. En segundo lugar, algunos niños no reciben una amplificación óptima porque necesitan ayudas más potentes o más apropiadas. En tercer lugar, muchos niños están utilizando audífonos que no están en buenas condiciones. Reichman y Healy (1989), examinaron el estado de monitorización y mantenimiento de sistemas de amplificación utilizados en colegios, y encontraron que el 54,1% de la muestra no había logrado llegar al nivel que asegurara que los audífonos estaban funcionando apropiadamente.

#### **2.3.1.7.6. Desarrollo de Técnicas de Comunicación**

Las tres principales variables del proceso global de comunicación en alumnos con deficiencias auditivas son, el grado de pérdida auditiva, nivel de inteligibilidad del habla, y modo de comunicación expresiva. La relación entre las tres es altamente significativa.

El uso de la comunicación manual, por parte de un alumno está muy relacionado con el hecho de que si el alumno atiende un programa educativo donde se utilizan los signos para su instrucción.

#### **2.3.1.7.7. Adaptar Evaluación y Entrenamiento**

La mayor parte de la evaluación y del entrenamiento descansa, en gran parte, en instrucciones e indicaciones verbales. Estas alternativas podrán requerir un aprendizaje considerable por parte de los padres, de los cuidadores y del resto de los miembros del equipo. Si se utiliza un lenguaje alternativo (lenguaje gestual), sus compañeros de comunicación deberán ser los suficientemente competentes en esos lenguajes de modo que puedan proporcionar un entorno de lenguaje más rico.

### **2.4. Hipótesis.**

¿El diseño de un sistema de codificación de señales de voz a texto usando Matlab mejorará la capacidad de comunicación de las personas con deficiencia auditiva?

**Variable Independiente:** Sistema de Codificación de señales de voz a texto usando Matlab.

**Variable Dependiente:** Capacidad de comunicación de Personas con Deficiencia Auditiva.

## **CAPÍTULO III**

### **3. MARCO METODOLÓGICO:**

#### **3.1.Enfoque.**

Por la naturaleza de las variables del problema de comunicación en personas con deficiencia auditiva; la investigación se efectuó con un análisis y un estudio analítico - descriptivo, por estar relacionado con técnicas de procesamiento digital de señales así como la determinación de una sistema adecuado para mejorar la comunicación de la personas con deficiencia auditiva.

#### **3.2.Modalidad básica de investigación.**

##### **3.2.1. Bibliográfico documental.**

Para poder desarrollar este proyecto necesariamente se realizó una investigación de tipo documental; debido que se acudió a textos de programación acerca del software Matlab, así como sobre electrónica especializada en procesamiento digital de señales y codificación de señales, los cuales nos dotaron de valiosa información técnica que será de gran ayuda para aplicar en nuestro trabajo.

### **3.2.2. De campo.**

Una vez adquirida la información teórica y técnica necesaria, se tuvo la necesidad de acudir al centro educativo para personas con sordera a verificar los métodos de comunicación y los problemas asociados a ellos, comprobando su limitación en el aprendizaje así como en el uso tecnológico y bibliográfico.

### **3.2.3. Aplicada**

Con la implementación del presente sistema se puso en práctica todo lo relacionado con la electrónica digital y en especial al análisis y procesamiento de señales.

## **3.3. Tipo o nivel de investigación.**

### **3.3.1. Exploratoria.**

El nivel de investigación fue exploratorio en cuanto se buscó datos documentados para contextualizar el problema planteado que se encuentra dentro del marco de indagación del presente proyecto, el mismo no ha sido investigado y en donde sus variables son de interés indagatorio.

### **3.3.2. Descriptiva.**

Descriptivo porque es una investigación de interés social, donde las personas con deficiencia auditiva van a poseer un sistema para resolver sus problemas de comunicación; además permitió evaluar y medir el grado de relación entre variables.

Existe asociación de variables puesto que analiza la vinculación de las personas discapacitadas con la necesidad de comunicación con el resto de la sociedad y de esta manera determinar la ejecución del proyecto.

### **3.4. Población y muestra.**

Las personas que formarán parte de la encuesta son el personal docente del Instituto de Educación Especial Dr. Camilo Gallegos, por facilidad de adquisición de información, ya que es imposible realizar la encuesta a las personas con discapacidad auditiva. .

El número de personas que forman parte de nuestra encuesta son 14, por tal motivo no será necesario realizar una muestra de población.

### **3.5. Métodos y técnicas de la investigación.**

**Encuesta.-** Se preparó una lista de preguntas para el personal docente del Instituto Educativo Dr. Camilo Gallegos para identificar los problemas con mayor facilidad y eficiencia.

### **3.6. Operacionalización de las variables.**

**Variable Independiente:** Sistema de Codificación de señales de voz a texto usando Matlab.

Tabla 3.1. Tabla de Operacionalización de la Variable Independiente  
Fuente: Investigador  
Elaborado por: Investigador

CONCEPTUALIZACIÓN	DIMENSIONES	INDICADORES	ITEMS	TÉCNICAS	INSTRUMENTOS
Es un método por el cual las señales de voz son procesadas y tratadas, para ser transformadas, por medio de algoritmos matemáticos, a texto.	Señales de voz  Procesadas  Transformada  algoritmos	El Habla  Receptor  Codificación  Programación	Función de Conversores A/D  Algoritmos de Codificación  Matlab	Observación	Guía de Observación

**Variable Dependiente:** Personas con Deficiencia Auditiva.

Tabla 3.2. Tabla de Operacionalización de la Variable Dependiente  
Fuente: Investigador

Conceptualización	Dimensión	Indicadores	Ítems	Instrumento
Es aquella que no puede escuchar normalmente debido a algún tipo de anomalía en el órgano de la audición:	Escuchar  Anormalidad  Órgano  Audición	Comunicación  Degeneración  Oído  Sentido	Cuál es el medio que utiliza para comunicarse con los alumnos con discapacidad auditiva?  Conoce usted si la discapacidad auditiva de los alumnos es degenerativa?  Conoce si existe algún alumno que antes pudo escuchar?  Cree usted que con la ayuda de la tecnología, los alumnos sordo-mudos puedan tener la	Cuestionario



			<p>oportunidad de desarrollar nuevos métodos de comunicación?</p> <p>Los alumnos conocen los símbolos alfanuméricos?</p> <p>Cree usted que con la implementación del software los alumnos podrán desenvolverse de mejor manera en la sociedad?</p> <p>¿Piensa usted que en un futuro las personas con deficiencia auditiva puedan desarrollar la habilidad de leer textos en español?</p>	
--	--	--	---	--

### 3.7. Plan de Recolección de Información

Tabla 3.3. Plan de Recolección de Información

Fuente: Investigador

<b>PREGUNTAS BÁSICAS</b>	<b>EXPLICACIÓN</b>
1.- ¿Para qué?	Para determinar si el sistema de codificación de voz a texto mejorará en gran volumen la comunicación entre personas con discapacidad auditiva y personas normales.
2.- ¿De qué personas?	Personas con deficiencia auditiva.
3.- ¿Sobre qué aspectos?	La comunicación entre personas normales y sordo-mudas
4.- ¿Quién?	Andrés Cárdenas.
5.- ¿A quiénes?	Personas con discapacidad Auditiva del Instituto Dr. Camilo Gallegos.
6.- ¿Cuándo?	6 meses a partir de la aprobación del proyecto
7.- ¿Dónde?	Ciudad de Ambato
8.- ¿Cuántas veces?	Una encuesta y una entrevista
9.- ¿Qué técnicas de recolección?	Encuesta, entrevista.
10.- ¿Con qué?	Cuestionario y lista de cotejo.

### **Plan de Recolección de la Información**

Se realizó una revisión crítica de información, donde los datos fueron estudiados punto por punto conforme a los criterios expresados por los encuestados, a su vez en los casos donde existieron problemas, necesariamente se repitió la recolección de datos.

Su análisis se efectuó mediante una representación gráfica por medio de tablas y figuras, así como escrita dando una explicación de los resultados de cada punto de la encuesta.

## CAPÍTULO IV

### 4. Análisis de resultados.

#### 1. ¿Cuál es el medio que utiliza para comunicarse con los alumnos con discapacidad auditiva?

Tabla 4.1. Resultados - Pregunta 1  
Elaborado por: Investigador

Ítem	Nº	Porcentaje
Lenguaje de señas	14	100%
Lectura de labios	0	0%
<b>Total</b>	<b>14</b>	<b>100%</b>



Gráfico 4.1 Resultados - Pregunta 1  
Elaborado por: Investigador

Según los datos obtenidos el principal método de comunicación utilizado para impartir clases en el Instituto Educativo es el lenguaje de señas, siendo el único método de comunicación utilizado dentro de la escuela.

**2. ¿Conoce usted si la discapacidad auditiva de los alumnos es degenerativa?**

Tabla 4.2. Resultados - Pregunta 2  
Elaborado por: Investigador

Ítem	Nº	Porcentaje
Si	8	57.14%
No	6	42.86%
<b>Total</b>	<b>14</b>	<b>100%</b>

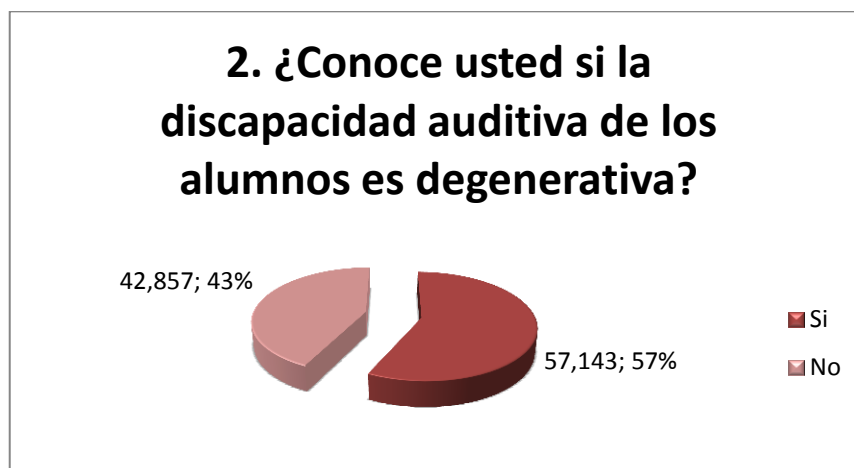


Grafico 4.2. Resultados - Pregunta 2  
Elaborado por: Investigador

El 57% de los profesores conocen las causas de la discapacidad de sus alumnos, donde la mayoría de los encuestados coinciden que la deficiencia auditiva es congénita, mientras el 43% asegura no conocer las causas de la discapacidad de los niños por tal razón no emiten un criterio afirmativo sobre la presente pregunta de la encuesta.

### 3. ¿Conoce usted si existen alumnos que antes podían escuchar?

Tabla 4.3. Resultados - Pregunta 3  
Elaborado por: Investigador

Ítem	Nº	Porcentaje
Si	5	35.71%
No	9	64.29%
<b>Total</b>	<b>14</b>	<b>100%</b>



Figura 4.3 Resultados - Pregunta 3  
Elaborado por: Investigador

El 35.71% de los encuestados conocen si los estudiantes en un principio podían escuchar, los cuales afirman que los niños en su totalidad nacieron con deficiencia auditiva, mientras el porcentaje restante asegura no saber si sus alumnos alguna vez pudieron escuchar.

Los datos obtenidos en la presente pregunta empatan con los resultados adquiridos en la pregunta anterior.

**4. ¿Cree usted que con la ayuda de la tecnología, los alumnos sordo-mudos puedan tener la oportunidad de desarrollar nuevos métodos de comunicación?**

Tabla 4.4. Resultados - Pregunta 4

Elaborado por: Investigador

Ítem	Nº	Porcentaje
Si	10	71.42%
No	4	28.58%
<b>Total</b>	<b>14</b>	<b>100%</b>

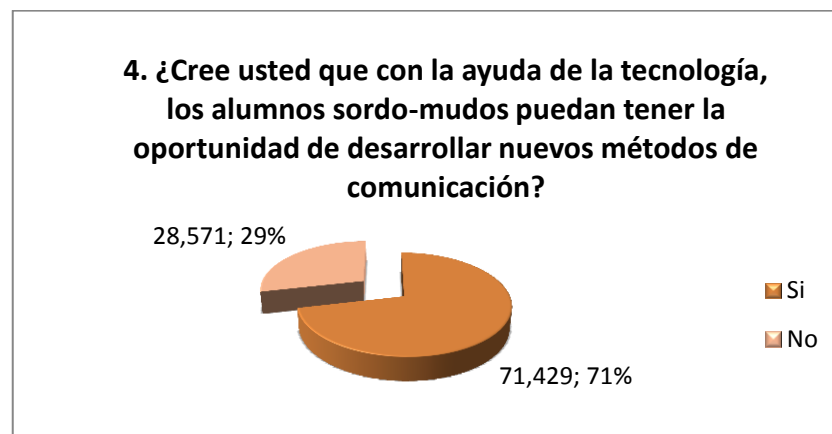


Gráfico 4.4 Resultados - Pregunta 4

Elaborado por: Investigador

El 71% de los encuestados piensan que con la evolución de la tecnología, las personas con deficiencia auditiva puedan desarrollar nuevas técnicas de comunicación así como involucrarse de mejor manera en la sociedad, frente a un 29% donde se muestran escépticos con la ayuda de la tecnología que pueda brindar a la discapacidad auditiva, por falta de interés en el desarrollo de aplicaciones para personas discapacitadas.

5. ¿Los alumnos conocen los símbolos alfanuméricos?

Tabla 4.5. Resultados - Pregunta 5  
Elaborado por: Investigador

Ítem	Nº	Porcentaje
Si	3	21,4%
No	11	78,6%
Total	14	100%

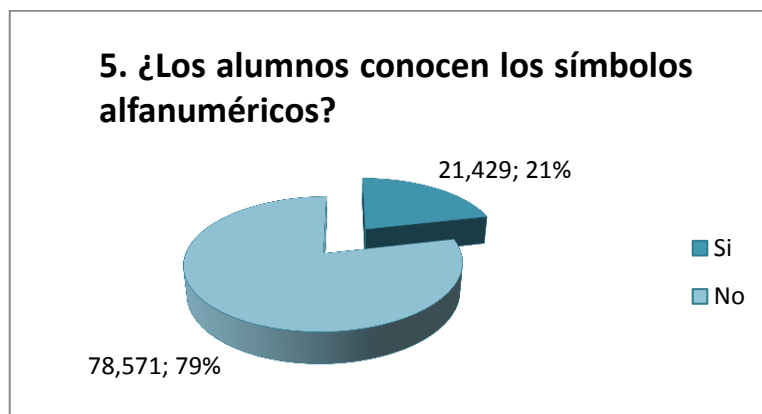


Figura 4.5. Resultados - Pregunta 5  
Elaborado por: Investigador

Del total de los docentes, el 79% aseguran la falta de conocimiento de sus alumnos sobre los símbolos alfa-numéricos, de forma contraria el 21% afirman el conocimiento superficial de sus alumnos sobre los símbolos, dicho en otras palabras los han visto en textos y computadoras pero no los saben reconocer.



**6. ¿Cree usted que con la implementación del software los alumnos podrán desenvolverse de mejor manera en la sociedad?**

Tabla 4.6. Resultados - Pregunta 6  
Elaborado por: Investigador

Ítem	Nº	Porcentaje
Si	9	64.29%
No	5	35.71%
<b>Total</b>	<b>14</b>	<b>100%</b>

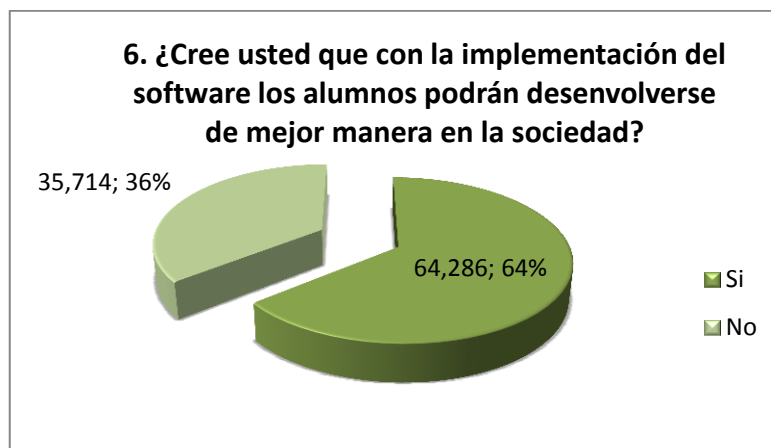


Gráfico 4.6. Resultados - Pregunta 6  
Elaborado por: Investigador

El 64% de los encuestados se muestran optimistas con la ayuda de un software en la impartición de clases, los alumnos podrán aprender a reconocer los símbolos del lenguaje español, el cuál les permitirá en un futuro desenvolverse de mejor manera dentro de la sociedad. Mientras que el 36% no ven la forma de cómo les podría ayudar el software a sus alumnos en aprender los símbolos alfa-numéricos utilizados en el lenguaje español y por ende su involucración normal en la sociedad.

**7. ¿Piensa usted que en un futuro las personas con deficiencia auditiva puedan desarrollar la habilidad de leer textos en español?**

Tabla 4.7. Resultados - Pregunta 7  
Elaborado por: Investigador

Ítem	Nº	Porcentaje
<b>si</b>	8	57%
<b>No</b>	6	43%
<b>Total</b>	<b>45</b>	<b>100%</b>



Gráfico 4.7. Resultados - Pregunta 7  
Elaborado por: Investigador

Existe una opinión un tanto pareja entre los docentes, frente a la posibilidad de desarrollar una habilidad en sus alumnos después de usar la aplicación, la cual les permitirá leer textos escritos en lenguaje español, pero sin embargo un poco más de la mitad de encuestados se muestran optimistas frente a la respuesta que puedan tener sus dirigidos al impartirles un nuevo método de comunicación.

## **CAPÍTULO V**

### **5. CONCLUSIONES Y RECOMENDACIONES**

#### **5.1. Conclusiones**

- Después de haber realizado la encuesta, los docentes creen factible la realización de una aplicación que permita transformar señales de voz a texto.
- Con la implementación del sistema, los alumnos tendrán a disposición un nuevo método de comunicación.
- Las personas con deficiencia auditiva tendrán la posibilidad de desenvolverse de mejor manera en la sociedad.

#### **5.2. Recomendaciones**

- Implementar un sistema de transformación de señales de voz a texto con una interfaz gráfica, desarrollada en guide de matlab, que sea de fácil manejo y entendimiento para los docentes y alumnos del centro educativo para discapacitados.
- Utilizar procesos de codificación de señales que permitan extraer señales puras con características fáciles de manipular.
- Realizar pruebas de transformación de señales de voz en distintos ambientes y provenientes de distintas personas

## **CAPÍTULO VI**

### **6. Propuesta.**

Sistema de codificación de señales de voz a texto usando Matlab para personas con deficiencia auditiva del Instituto de Educación Especial Dr. Camilo Gallegos.

#### **6.1. Datos informativos.**

Institución ejecutora: Instituto de Educación Especial Dr. Camilo Gallegos

Ubicación: Ciudadela La Pradera

Beneficiarios: Estudiantes con deficiencia auditiva del Instituto Dr. Camilo Gallegos

#### **6.2. Antecedentes de la propuesta**

Durante las observaciones realizadas en la escuela, se notó en diferentes clases dictadas el uso de lenguaje de señas como principal método de comunicación, donde todos los alumnos están capacitados para comprender dicho lenguaje. Pero cuando se necesitaba impartir ciertos temas donde los alumnos debían usar textos escritos con el alfabeto, se observó la necesidad de buscar una forma que los estudiantes aprendan a reconocer los símbolos del alfabeto.

Por otro lado, la deficiencia auditiva siempre ha sido un obstáculo para las personas que lo padecen, porque los limita hasta cierto punto al momento de realizar diferentes tipos de trabajos y en ciertas actividades diarias de la vida. Con el transformador de señales de voz a texto como herramienta de aprendizaje, a las personas discapacitadas, se las está brindando o poniendo a disponibilidad un nuevo método de comunicación, el cual se espera que en un futuro, sea su instrumento diario de interacción con la sociedad en general.

Por tanto se espera el desarrollo del sistema les permita familiarizarse de una manera fácil y dinámica con los símbolos alfanuméricos del lenguaje español, en donde los profesores junto con el sistema juegan un papel fundamental en el proceso de enseñanza de una nueva forma de comunicación, para los niños con deficiencia auditiva.

### **6.3. Justificación**

A través del presente proyecto los estudiantes con deficiencia auditiva del Instituto Educativo Dr. Camilo Gallegos logran desarrollar una nueva forma de comunicación, de una u otra forma los ayuda a involucrarse de mejor manera a la sociedad, de tal manera pueden responder con mayor solvencia en todos los ámbitos sociales.

El programa se orienta a escolares, con el fin de familiarizarlos desde cortas edades con los símbolos utilizados por el lenguaje español. Ésta aplicación será de gran ayuda por ayudar a familiarizar los símbolos alfanuméricos con los del lenguaje de señas que las personas sordo-mudas lo manejan.

Los principales beneficiados como se explicó anteriormente, son los alumnos con deficiencia auditiva por tener mayor interacción con la sociedad. Además se mejora su comunicación con personas normales que no manejen el lenguaje de señas, por tal razón el Instituto Educativo Dr. Camilo Gallegos podrá impartir

mayor cantidad de materias donde se utilice el abecedario, así los estudiantes pueden desenvolverse de mejor manera en clases, y en un futuro puedan incluso leer textos escritos en lenguaje español.

#### **6.4. Objetivos.**

##### **6.4.1. Objetivo general.**

Implementar un sistema de codificación de señales de voz a texto, para personas con deficiencia auditiva del Instituto de Educación Especial Dr. Camilo Gallegos usando Guide de Matlab.

##### **6.4.2. Objetivos Específicos.**

- Explicar los métodos de procesamiento o tratamiento digital de una señal de voz mediante filtros, normalización de datos y transformada rápida de Fourier (FFT),
- Diseñar una aplicación que permita codificar señales de voz a texto utilizando la interfaz Guide de Matlab.
- Realizar pruebas de reconocimiento y transformación de voz a texto con la aplicación desarrollada.

## **6.5. Análisis de factibilidad**

Este proyecto es factible por cuanto su desarrollo no implica el uso de equipos o circuitos externos, reduciendo el gasto de dinero en el desarrollo. Cabe recalcar que todos los egresos generados en el proceso de investigación y desarrollo, fueron por parte del investigador, el Instituto no se comprometió a ayudar económicamente con el proyecto.

La parte investigativa se tornó viable, en la biblioteca de la F.I.S.E.I. existe material bibliográfico acorde a las necesidades, los cuales fueron de gran ayuda como material de consulta. A nivel del internet (web) de la misma forma existe gran diversidad de archivos y libros digitales relacionados con Procesamiento Digital de Señales, Codificación Digital de Señales, los mismos fueron la base fundamental en el proceso de investigación técnica.

El medio por donde se desarrolló el sistema de transformación de señales, es una herramienta muy potente de cálculo, la cual nos facilitó al momento de efectuar los procesos de cálculos matemáticos y eliminándonos la necesidad de fabricar o comprar circuitos o elementos electrónicos para implementar. Debido a sus diversas funciones reemplaza los procesos de algunos elementos electrónicos.

## **6.6. Fundamentación**

### **Codificación Digital de Señales**

Existe un amplio conjunto de sistemas donde son de especial importancia algoritmos sofisticados y hardware para tratamiento de señales, sistemas que van desde sistemas militares altamente especializados, pasando por aplicaciones industriales, hasta llegar a la electrónica de consumo de bajo coste y altos volúmenes de ventas.

El tratamiento de señales trata de la representación, transformación y manipulación de señales y de la información que contienen.

Antes de los años 60, la tecnología para el tratamiento de señales era casi exclusivamente analógica en tiempo continuo. La rápida evolución de los computadores y los microprocesadores digitales junto con algunos importantes desarrollos fueron la causa de un importante desplazamiento hacia las tecnologías digitales, dando lugar al campo del tratamiento digital de señales.

El tratamiento de señales digitales se basa en el procesamiento de secuencias de muestras.

En tratamiento digital de señales, las señales se representan como secuencias de números de precisión finita y el procesado se realiza utilizando cómputo digital. El término más general "tratamiento de señales en tiempo discreto" incluye el tratamiento digital de señales como un caso especial, pero incluye también la posibilidad de que las secuencias de muestras sean procesadas con otras tecnologías en tiempo discreto.

La mayoría de las aplicaciones requieren el uso de tecnología de tiempo discreto para procesar señales en tiempo continuo. En este caso, las señales en tiempo continuo se convierten en una secuencia de muestras, es decir, en señales en tiempo discreto. Tras este proceso en tiempo discreto las señales de salida se vuelven a convertir en señales en tiempo continuo. En general buscamos que este



proceso se de en tiempo real, lo que significa que el sistema en tiempo discreto se implementa de forma que las muestras de la salida se calculan a la misma velocidad a la que se muestrea la señal en tiempo continuo.

### 6.6.1. Transformada de Fourier

La transformada de Fourier, en esencia, descompone o expande una señal o función en senos y cosenos de diferentes frecuencias cuya suma corresponde a la señal original, es decir, es capaz de distinguir las diferentes componentes de frecuencia de la señal, y sus respectivas amplitudes. La transformada de Fourier de una función del tiempo  $f(t)$  se define como:

Ec. 5

$$F(w) = \int_{-\infty}^{\infty} f(t)e^{iwt} dt$$

De acuerdo con lo dicho anteriormente la transformada de Fourier puede obtener un representación en el dominio de la frecuencia de una señal que se encuentra originalmente en el dominio del tiempo.

#### 6.6.1.1. Transformada discreta de Fourier (DFT)

“La transformada discreta de Fourier (*DFT* por sus siglas en inglés) permite evaluar la transformada de Fourier de secuencias de duración finita. La *DTF* es una secuencia compleja que es obtenida por medio de muestrear un período de la transformada de Fourier de la señal a un número finito de puntos de frecuencia, es decir, que corresponde a muestras igualmente espaciadas en frecuencia de la transformada de Fourier de la señal discreta”<sup>13</sup>. La *DFT* es importante por dos razones. Primero, permite determinar el contenido frecuencial de la señal de voz, o sea, realizar análisis espectral. La segunda razón de importancia es realizar

---

<sup>13</sup>Borrás Gené Oriol, *Reductor de ruido mediante resta espectral en entorno Matlab, 2006*

operaciones de filtrado en el dominio de la frecuencia. La eficiencia es la razón principal por la cual se procesan las señales en el dominio de la frecuencia.

### Muestreo de la transformada de Fourier<sup>14</sup>

Consideremos una secuencia periódica  $x[n]$  con su transformada de Fourier  $X(e^{j\omega})$  y asumamos que una secuencia  $X[k]$  es obtenida al muestrear  $X(e^{j\omega})$  a frecuencias  $\omega_k = \frac{2\pi k}{N}$ , como sigue:

**Ec. 6**

$$X[k] = X(e^{j\omega}) \Big|_{\omega = \frac{2\pi k}{N}} = X\left(e^{j\left(\frac{2\pi k}{N}\right)}\right)$$

Como la transformada de Fourier es periódica en  $\omega$  con período  $2\pi$ , la resultante secuencia es periódica en  $k$  con período  $N$ . La secuencia de muestras es periódica dado a que los  $N$  puntos están igualmente espaciados iniciando desde cero. Por lo que la misma secuencia se repite mientras  $k$  varía en el rango de  $0 \leq k \leq N - 1$ .

Se puede notar que la secuencia de muestras  $X[k]$ , siendo periódica con período  $N$ , podría ser la secuencia de los coeficientes discretos de la serie de Fourier de la secuencia  $x^{\wedge}[n]$ , la cual tiene la relación con  $x[n]$  del siguiente modo:

**Ec. 7**

$$x^{\wedge}[n] = \sum_{r=-\infty}^{\infty} x[n - rN] = x[n] * \sum_{r=-\infty}^{\infty} \delta[n - rN]$$

Esto es  $x^{\wedge}[n]$  es la secuencia periódica que resulta de la convolución aperiódica de  $x[n]$  con un tren periódico de impulsos unitarios. Por tanto, la secuencia periódica  $x^{\wedge}[n]$ , correspondiente a  $X[k]$ , que es obtenida de muestrear  $X(e^{j\omega})$ , esta formada

---

<sup>14</sup> Borrás Gené Oriol, *Reductor de ruido mediante resta espectral en entorno Matlab*, 2006

de  $x[n]$  por medio de sumar juntos un número infinito de replicas cambiadas de  $x[n]$ . Estos cambios son todos los enteros positivos y negativos múltiplos de  $N$ . El período de la secuencia  $X[k]$ .

Equivalentemente,  $x[n]$  se puede recuperar a partir de la correspondiente secuencia periódica  $x^{\wedge}[n]$  a través de la siguiente ecuación:

**Ec. 8**

$$x[n] = \begin{cases} x^{\wedge}[n], & 0 \leq n \leq N - 1 \\ 0, & \text{otherwise.} \end{cases}$$

Alternativamente, dada la secuencia de los coeficientes de Fourier  $X[k]$ , se puede encontrar  $x^{\wedge}[n]$  y utilizar la ecuación anterior para obtener  $x[n]$ . Cuando las series de Fourier son utilizadas de este modo para representar secuencias de duración finita, es llamada la transformada discreta de Fourier.

### 6.6.1.2. Transformada Rápida de Fourier

La transformada rápida de Fourier tiene gran importancia en una gran variedad de aplicaciones, como ejemplo el procesamiento digital de señales.

FFT es la abreviatura usual (de sus siglas en inglés Fast Fourier Transform), y es un eficiente algoritmo que permite calcular la transformada discreta de Fourier y su inversa dados vectores de longitud  $N$  por las siguientes ecuaciones:

**Ec. 9**

$$X(k) = \sum_{j=1}^N x(j) w_N^{(j-1)(k-1)}$$

## Ec. 10

$$x(j) = \left(\frac{1}{N}\right) \sum_{k=1}^N X(k) w_N^{-(j-1)(k-1)}$$

Donde:

$$w_N = e^{\frac{-2\pi i}{N}}$$

Es una N-ésima raíz de unidad.

### 6.6.2. Filtrado de Señales

Casi todos los sistemas de comunicación emplean filtros. Un filtro deja pasar una banda de frecuencia mientras rechaza otras. Los filtros pueden separar las señales deseadas de las no deseadas.

“La teoría de filtros tiene una componente matemática compleja que hace que se aleje de la experiencia humana. La ecuación de un filtro, por ejemplo, no está relacionada necesariamente con sus cualidades sonoras. En los textos técnicos, los filtros se describen mediante una herramienta matemática denominada transformada z.”<sup>15</sup>

La transformada z es un concepto esencial para el diseño de filtros, ya que proporciona una relación matemática entre las características del filtro que queremos diseñar y los parámetros de implementación del mismo.

#### 6.6.2.1. Tipos de filtros

Los filtros más corrientes son los filtros paso bajo (Low Pass, LP), paso alto (High Pass, HP), paso de banda (Band Pass, BP) y los filtros rechazo de banda (o paso

---

<sup>15</sup> Borrás Gené Oriol, *Reductor de ruido mediante resta espectral en entorno Matlab*, 2006

no banda) (Band Reject, Band stop o Notch). En la figura 1 se representan estos 4 tipos de filtros mediante su respuesta en frecuencia o espectro de amplitud.

Cada punto de la respuesta en frecuencia nos indica la atenuación a la que se someterá una señal a una frecuencia determinada.

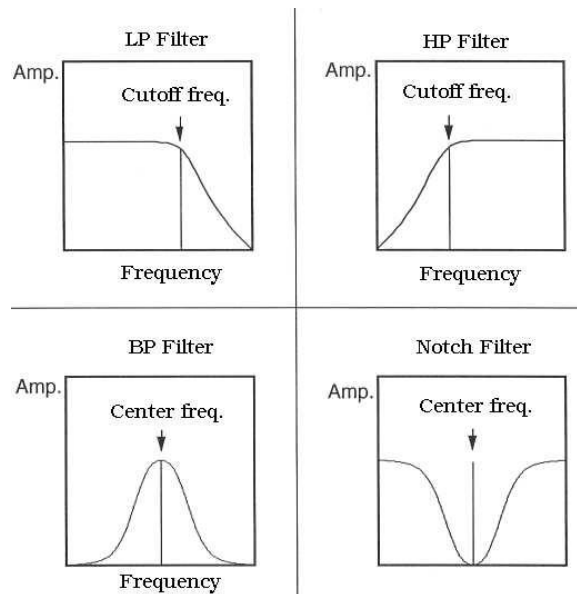


Figura 6.1. Tipos de Filtros

Fuente: Velásquez Ramírez Genoveva, *Sistema de Reconocimiento de Voz*

- Los **filtros paso bajo (LP)** dejan pasar las frecuencias que están por debajo de una determinada frecuencia.
- Los **filtros paso alto (HP)** dejan pasar las frecuencias que están por encima de una determinada frecuencia.

Estos dos tipos de filtros están definidos por su **frecuencia de corte**, que es la frecuencia a la cual la amplitud de la señal se reduce a  $0.707 \left(\frac{1}{\sqrt{2}}\right)$  de su valor máximo, es decir, sufre 3 dB de atenuación.

- Los **filtros paso banda (BP)** dejan pasar las frecuencias que están situadas en una determinada banda de frecuencia, es decir, entre dos determinadas frecuencias.
- Los **filtros rechazo de banda (BR)** dejan pasar todas las frecuencias excepto las que están situadas en una determinada banda de frecuencia, es decir, entre dos determinadas frecuencias  $f_1$  y  $f_2$ . Estas frecuencias son las frecuencias a las que la amplitud de la señal se reduce a  $0.707 \left(\frac{1}{\sqrt{2}}\right)$  de su valor máximo, es decir, sufre 3 dB de atenuación.

Estos dos tipos de filtros están definidos por su frecuencia central y su ancho de banda, que sería la diferencia entre las frecuencias de corte inferior y superior. Los filtros pueden combinarse en serie o en paralelo para obtener respuestas frecuenciales más complejas.

### **6.6.3. Evaluación del Sistema de Transformación de Señales de Voz**

Para evaluar la calidad del sistema de transformación después de realizar el proceso de conversión de señales de voz a texto, a partir de la interfaz diseñada en Matlab, se efectuó una serie de transformaciones por distintas personas y expresando diferentes palabras en distintos ambientes, los cuales nos propinaron distintos datos que son analizados en los siguientes puntos.

### **6.6.4. Parámetros para evaluar la calidad de transformar la señales de voz**

Para analizar la implementación del programa, a todo el grupo de señales introducidas, se debe tener en cuenta los siguientes parámetros.

1. Para las señales de voz:
  - Acentuación, niveles, volumen de voz
2. Condiciones de la emisión del habla:
  - Ambiente ideal (silencioso), ruido moderado, y ruidoso (alto ruido)

## 6.7. Metodología

A continuación se muestra el diagrama propuesto, el cual consiste de tres etapas: la etapa de la captura de la señal de voz mediante micrófono, la etapa de procesamiento de la señal y finalmente la etapa de verificación del hablante empleando las características extraídas en las dos primeras etapas.

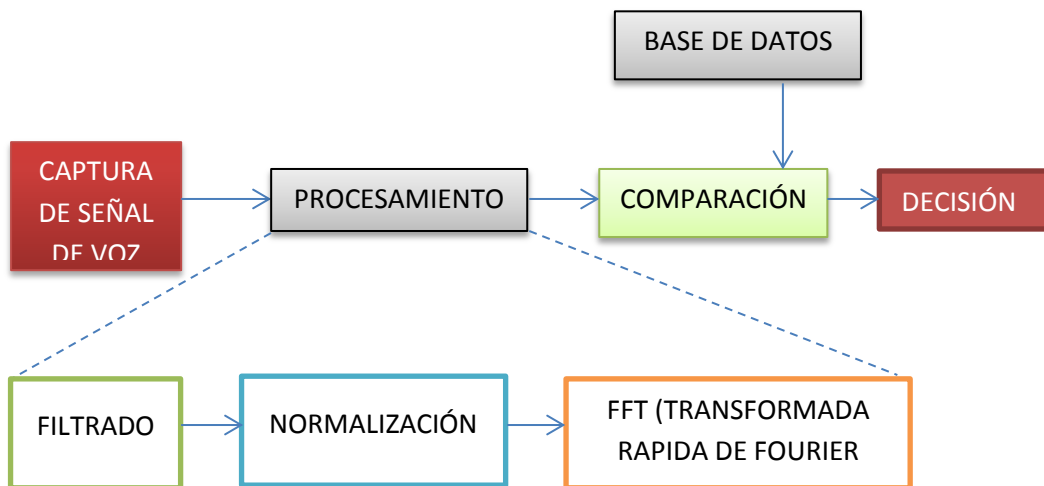


Figura 6.2 Diagrama de Bloques de un sistema general de conversión de señales de voz a texto

Elaborado por: Investigador

### 6.7.1. Proceso de Adquisición de la Señal del Hablante

En casi cualquier computadora personal se tiene disponible una tarjeta de sonido. También se pueden encontrar programas de cómputo para grabar y reproducir señales de audio. Usualmente, dichos programas vienen incluidos en Windows, Linux, y demás sistemas operativos. Algunos sistemas multimedia incluyen un programa para desplegar y editar estas señales de audio usando diferentes formatos estándar como: wav, VOC, mp3, entre otros.

Estos programas de cómputo pueden incluir diferentes opciones de despliegue en el dominio de tiempo, o en el dominio de la frecuencia. La tarjeta de sonido proveerá la circuitería (hardware) necesaria para la adquisición de datos, usualmente con una resolución de 16 bits, con entrada/salida mono/estéreo.

Para completar los elementos de trabajo, la tarjeta de adquisición de sonido debe ser complementada con el software Matlab que proveerá el ambiente de desarrollo para los algoritmos de Procesamiento Digital de Señales.

Cabe resaltar que se empleó una velocidad de muestreo de sonido de 11 KHz. Con la finalidad de cumplir con el criterio de Nyquist que es mayor ó igual a 2 veces la frecuencia de muestreo, que para nuestro caso pertenece al canal telefónico que es aproximadamente 4 KHz.

Tabla 6.1 Parámetros de la Tarjeta de Sonido

Elaborado por: Investigador

Velocidad de Transmisión	128 Kbps.
Tamaño de muestreo de sonido	16 bits.
Tipo de canal	Monofónico.
Velocidad de muestreo de sonido	11 KHz.
Formato de sonido	*.wav

## 6.7.2. Proceso de Filtrado de la Señal

### 6.7.2.1. Filtro Pasa Altas

Este tipo de filtro elimina todas las frecuencias desde cero hasta la frecuencia de corte y permite el paso de todas las frecuencias por encima de la frecuencia de corte. Con un filtro pasa alto, las frecuencias entre cero y la frecuencia de corte son la banda eliminada. Las frecuencias por encima de la de corte son la banda



pasante. Un filtro ideal pasa alto tiene una atenuación infinita en la banda eliminada, atenuación cero en la banda pasante y una transición vertical.

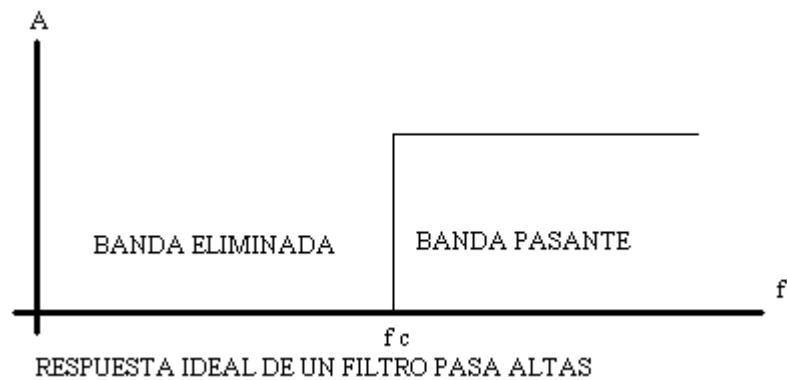


Figura 6.3. Respuesta de un Filtro Pasa Altas

Fuente: Velásquez Ramírez Genoveva, (2008), *Sistema de Reconocimiento de Voz*, Universidad San Carlos de Guatemala

#### 6.7.2.2. Filtro de Pre-Énfasis

El preénfasis acentúa las frecuencias altas de la señal de voz, esto debido a que el modelo del tracto vocal utilizado no filtra de buena manera las señales de frecuencias altas (no sonoras, por ejemplo: consonantes, “s”), a diferencia de las frecuencias bajas (sonoras, por ejemplo: vocales “a”).

Se aplica un filtro digital pasa altas de primer orden a la señal, para enfatizar las frecuencias altas de los formantes por dos razones, primero para que no se pierda información durante el procesamiento, ya que la mayoría de la información está contenida en las frecuencias bajas, en segundo remueve la componente DC de la señal, aplanando espectralmente la señal. Uno de los filtros de pre-énfasis más utilizados tiene la ecuación:

**Ec. 11**

$$H(z) = (1 - az^{-1})$$

$a = 0.95$  en nuestro caso.

Este filtro permitirá la eliminación de características no deseadas en la señal.

**6.7.3. Normalización**

Éste proceso consiste en ajustar todos los parámetros a una sola escala para que al momento de ser utilizados por el siguiente proceso (FFT), no causen problemas de estabilidad, en este caso la escala empleada se encuentra dada por los valores de  $[-1,1]$ , por lo tanto todos los archivos de la base de datos y de señales a ser reconocidas son normalizados a esta escala, y se encuentran almacenados dentro de un vector  $x[i]$ , con  $i=1,\dots,n$ .

Para lograr éste proceso se calcula el máximo y el mínimo del vector  $x[n]$ , se divide por el de mayor valor absoluto y todos los datos normalizados se encuentran dentro del intervalo  $[-1,1]$ .

**6.7.4. Transformada de la señal**

La FFT es un algoritmo que convierte una función compleja, muestreada en tiempo, en una función muestreada en dominio de la frecuencia y que también es compleja. Como la mayor parte de las veces se opera con valores reales y el algoritmo de la FFT requiere que la señal discretizada esté expresada en valores complejos, resulta imprescindible entonces utilizar la propiedad de que un número real puede ser expresado como un número complejo, si como valor imaginario le colocamos el valor cero. Ejemplo: el valor real 750 puede ser expresado como el número complejo  $(750, 0)$ , donde 750 es la parte real y 0 la parte imaginaria de ese número complejo.

Para aplicar la transformada rápida de Fourier hay que tomar en cuenta ciertos criterios:

- Se recomienda que la longitud del vector sea una potencia de 2.
- No se recomienda que la longitud del vector sea un número primo.
- Si la longitud del vector es menor que N, éste se rellena con ceros. Si es mayor, el vector es truncado.

Estos criterios son recomendaciones, mas no parte primordial al momento de realizar la FFT, pero para obtener mejores resultados se los puede aplicar.

#### **6.7.5. Diseño del sistema de codificación de señales de voz a texto utilizando la interfaz guide de Matlab.**

El objetivo del presente documento es definir de manera clara y precisa los procesos y funcionalidades del sistema desarrollado. En esta sección se presenta una descripción general del sistema además de los procesos para su diseño, con el fin de conocer las principales funciones que debe realizar, sus datos asociados, las restricciones impuestas, y cualquier factor que pueda afectar al desarrollo del mismo.

##### **6.7.5.1. Diseño del Sistema**

La interfaz del sistema de codificación de señales de voz a texto, fue desarrollado en guide de Matlab, la cual permite tener una idea más clara de cada uno de los proceso realizados en la transformación. La aplicación es bastante sencilla y amigable para el usuario. (Figura 6.4). (El flujograma y el código del programa se presentan en los Anexos B y C, respectivamente).

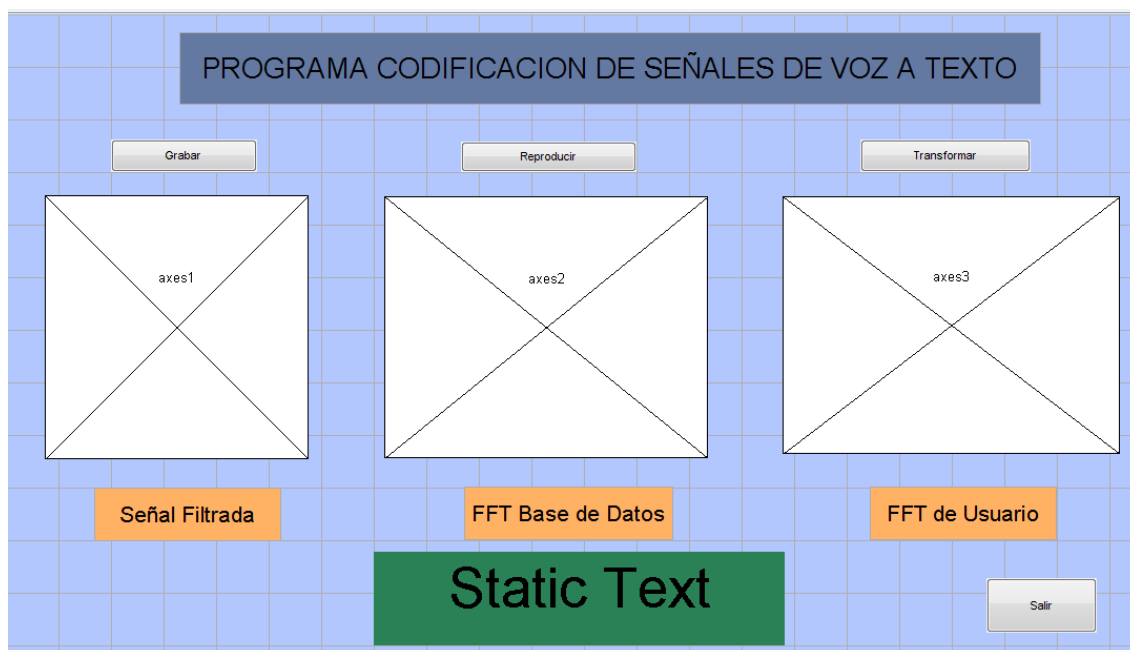


Figura 6.4. Interfaz Gráfica del Sistema de Transformación

Elaborado por: Investigador

Para realizar el proceso de transformación de señales de voz, primero, debemos asegurarnos que nuestra computadora, en donde vamos a usar el programa, tenga un micrófono correctamente conectado e instalado. Para constatar podemos acceder a panel de control, nos dirigimos a la pestaña de sonido, le damos clic y se despliega una ventana como se muestra en la Figura 6.5. Escogemos la pestaña **GRABAR** y observamos si nuestra PC tiene correctamente instalado un micrófono, lo probamos, y nos podemos dirigir a nuestra aplicación de codificación.

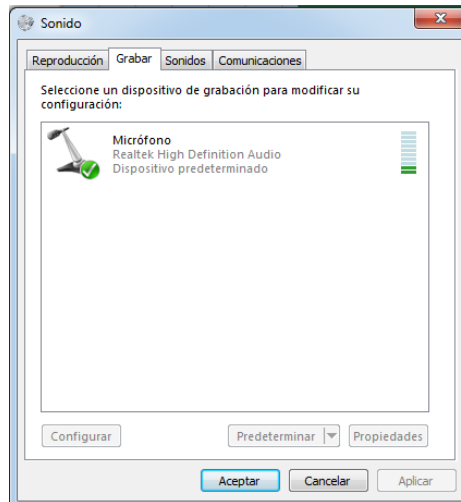


Figura 6.5 Vista de Configuración de Sonido

Elaborado por: Investigador

Una vez realizado el paso anterior, podemos ejecutar nuestro sistema de codificación. Para realizar la transformación de la señal de voz, en primer lugar obtenemos y grabamos una palabra emitida por una persona, esto mediante el comando **wavrecord**, para luego continuar con el proceso de filtrado de señal. La imagen de una señal capturada por una persona masculina, pronunciando el fonema “a” se muestra a continuación en la Figura 6.6.

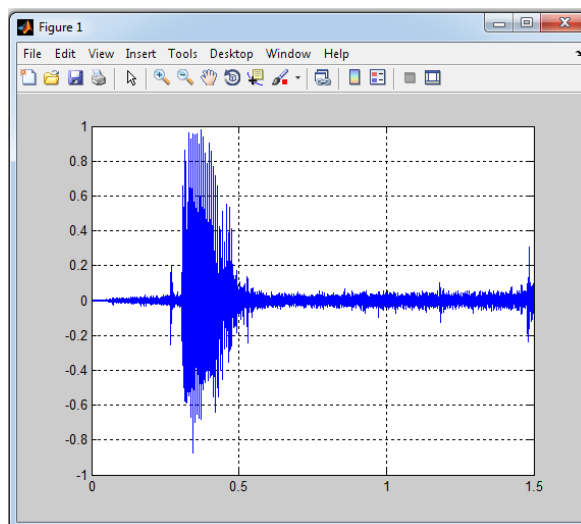


Figura 6.6 Señal del Fonema “a” en el Dominio del Tiempo

Elaborado por: Investigador

Como podemos observar la voz fue capturada en un intervalo de tiempo de 1,5 seg. A continuación la señal obtenida se la aplica un filtro de pre-énfasis pasa alto, que nos ayuda a enfatizar las altas frecuencias y por ende nos elimina las señales contenidas en bajas frecuencias. Para obtener el filtrado de la señal en Matlab utilizamos el comando **filter** y el resultado obtenido se muestra en la Figura 6.7.

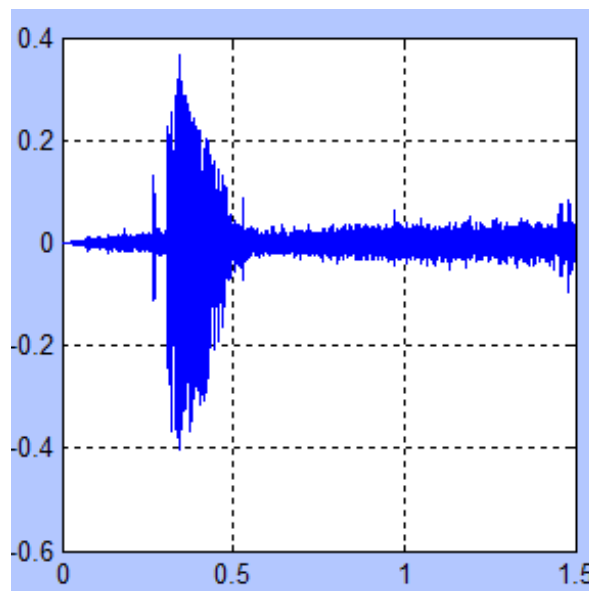


Figura 6.7 Señal Filtrada

Elaborado por: Investigador

Podemos observar claramente que en la señal filtrada de la Fig. se reduce el ruido y se disminuye la amplitud de la señal, quedándonos solo la energía y frecuencias necesarias para los siguientes procesamientos.

#### 6.7.5.2. Normalización de Señales

Como se puede observar la señal es muy variante y posee datos muy dispersos, lo cual, nos dificulta procesarla, por tanto es necesario ajustar dichos datos dentro de un rango específico para poder obtener características especiales. Para desarrollar dicho proceso aplicamos el siguiente proceso matemático.

```
maximo=max(abs(sonido));  
  
n=length(sonido);  
  
sonidoN=zeros(1,n);  
  
for i=1:1:n  
    sonidoN(i)=sonido(i)/maximo;  
end
```

Figura 6.8 Algoritmo del Proceso de Normalización

Elaborado por: Investigador

Dónde:

abs → retorna valores absolutos y magnitudes de números complejos

max → calcula los valores máximos a lo largo de las dimensiones de una matriz

length → calcula el tamaño máximo de las dimensiones de una matriz

zeros → crea una matriz de ceros

### 6.7.5.3. Cálculo de la FFT de la señal

La transformada rápida de Fourier nos permite transformar al dominio de la frecuencia una señal en el dominio del tiempo, pero como dicha transformada nos arroja datos positivos y negativos, le aplicamos la función de **abs** para obtener solo datos positivos y reales. Para obtener la transformada aplicamos la función **fft** de Matlab.

En la Figura 6.9, podemos observar la transformada rápida de Fourier del fonema “a”.

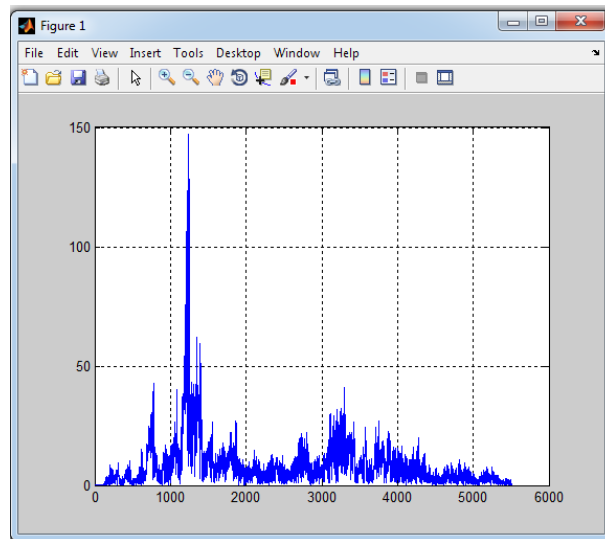


Figura 6.9 Señal en el Dominio de la Frecuencia (FFT)

Elaborado por: Investigador

Se observa claramente que el ancho de banda del fonema “a” se encuentra entre 1000 y 1400 Hz.

Como se explicó en los capítulos anteriores las frecuencias de una formante vocálica “a” se encuentran entre los rangos 800 a 1200 Hz. Estos valores pueden tener pequeñas variaciones debido a la forma de pronunciación, tipo de acento y tono de voz.

#### 6.7.5.4. Cálculo de proceso de Error

El proceso de cálculo de error nos ayuda a crear un vector con valores positivos y normalizados resultados u obtenidos de la resta entre las matrices de la fft de la señal guardada como base de datos y la fft de la señal ingresada.

Con la línea de código a continuación podemos realizar dicho proceso:

```
error(1)=mean(abs(transvocala-transusuario));
```

Una vez que se ha obtenido el vector error procedemos a calcular su valor mínimo con el comando **min**, para posteriormente hacer una sucesión de comparaciones



que nos permitirá determinar que palabra o fonema expresó el usuario, y desplegarlo en forma de texto.

```
error =  
      8.4524   9.5899   11.3036   8.5272   10.0921   10.1711   9.5770  
  
min_error =  
      8.4524
```

Figura 6.10 Resultados del Proceso de Error de la vocal “a”

Elaborado por: Investigador

En la Figura 6.10, podemos observar el vector **error** con todos sus valores medios y a continuación su **mínimo error**, que con las sucesivas comparaciones se puede constatar que el mínimo error es igual al valor en la primera posición del vector error. En consecuencia con dicha conclusión se puede tomar cualquier tipo de decisión, en nuestro caso desplegar en forma de texto la palabra pronunciada por el usuario.



Figura 6.11 Vocal Reconocida por el Sistema de Transformación

Elaborado por: Investigador

Continuando con el análisis del fonema “a”, una vez realizados todos los procesos ya explicados anteriormente, podemos obtener el fonema expresado por el usuario y visualizado como texto, tal como se muestra en la Figura 6.11.

#### **6.7.6. Modo de Uso**

En el presente punto, el objetivo es dar guía de uso de nuestra aplicación desarrollada, para que pueda ser manipulada con mayor facilidad en la escuela Dr. Camilo Gallegos y que sirva de una nueva herramienta de comunicación y enseñanza para los alumnos y profesores.

Hay que resaltar que no se ha desarrollado un archivo ejecutable del programa, por lo que se tendrá que instalar matlab en las computadoras que se necesite utilizar como muestra la aplicación desarrollada.

##### **6.7.6.1.Requisitos de Hardware y software**

**Hardware:** Se necesita una computadora con los siguientes requisitos mínimos.

- Procesador Pentium o AMD desde 1.33 GHz en adelante.
- Memoria RAM de 512 MB.
- Tarjeta de sonido con puertos de I/O de audio y entrada de MIC
- MIC en caso de usar una PC de tipo Desktop

**Software:** La PC debe tener instalado el siguiente software.

- Windows XP/Vista/Seven (7)
- Drivers de la tarjeta de sonido
- Matlab Versión 7.1.0.246 (R14) / 7.6.0.324 (R2008a)

### 6.7.6.2.Pasos para el uso de la aplicación

1. Ejecutar el programa Matlab que se ha instalado en la PC a usar. Se desplegará una ventana como se muestra a continuación en la Figura 6.12.

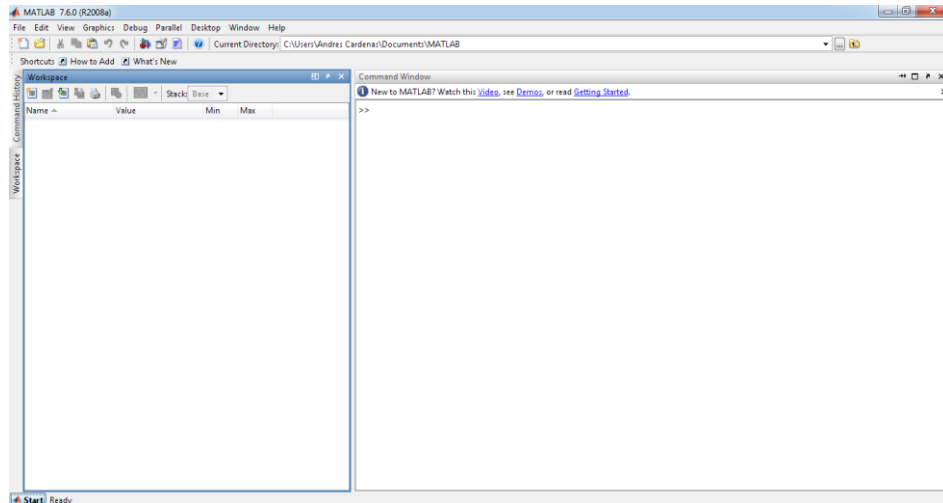


Figura 6.12 Pantalla Principal de Matlab

Elaborado por: Investigador

2. Para **abrir** la aplicación, se tiene 2 formas muy viables y sencillas. La primera dirigirse hacia la barra de herramientas, haciendo clic en la pestaña **File** como se muestra en la siguiente Figura 6.13.



Figura 6.13 Barra de Herramientas de Matlab

Elaborado por: Investigador

Se escoge nuestra aplicación con extensión **.fig**, para desplegar directamente nuestro programa.

La segunda forma es dirigirse a la subventana **Current Directory**, que se encuentra al lado izquierdo de la ventana de comandos, ahí podemos

seleccionar nuestra dirección donde tenemos alojados nuestros ficheros relacionados con nuestra aplicación, Figura 6.14.



Figura 6.14 Barra de Herramientas de Subventana de Matlab

Elaborado por: Investigador

Una vez seleccionado el **directorio**, en la ventana de comando digitamos el nombre de nuestra aplicación tal y como está redactada, en nuestro caso **TransformarVoz**, ya que Matlab diferencia entre mayúsculas y minúsculas. Ver Figura 6.15.

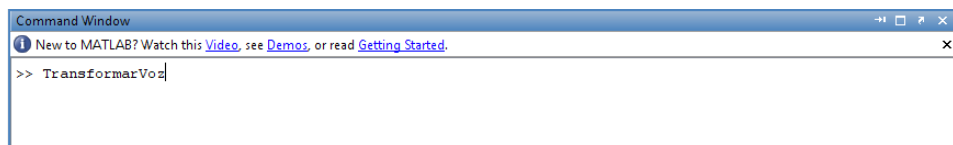


Figura 6.15 Ventana de Comandos de Matlab

Elaborado por: Investigador

3. Cuando se haya realizado uno de los dos pasos anteriores, se desplegará nuestro programa. Figura 6.16.

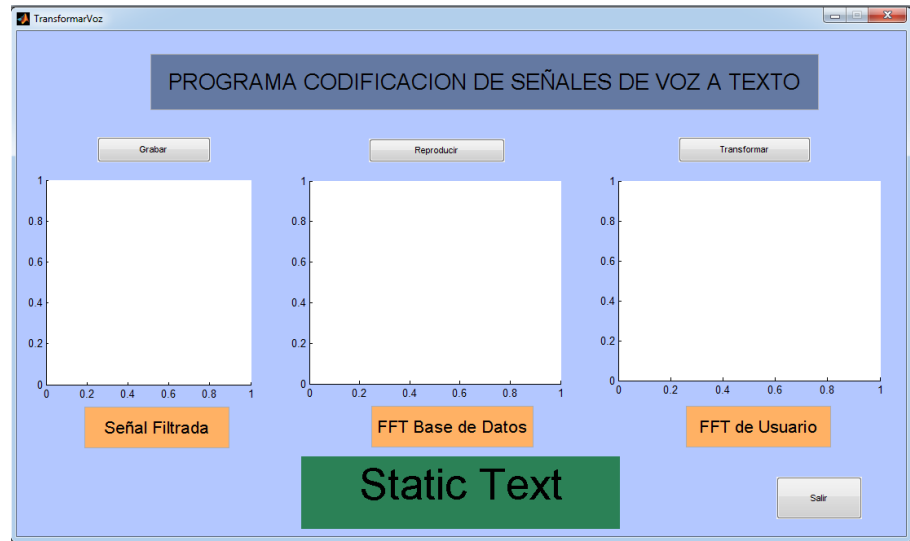


Figura 6.16 Pantalla Principal del Sistema de Transformación

Elaborado por: Investigador

4. Para poder realizar el proceso de transformar una señal de a voz a texto, se da clic en el botón **Grabar**. Figura 6.17.

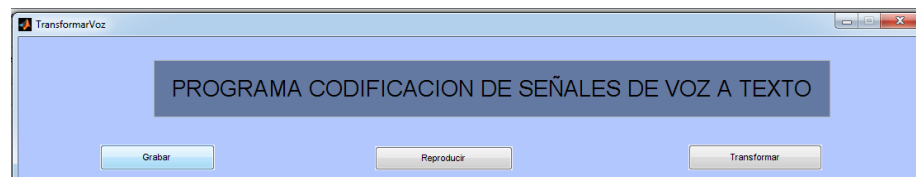


Figura 6.17 Sección de Control del Sistema de Transformación

Elaborado por: Investigador

5. Una vez expresada la palabra a ser reconocida, se grafica la señal introducida en el dominio del tiempo. Figura 6.18.

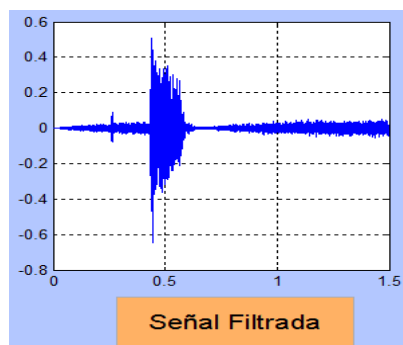


Figura 6.18 Señal Introducida Graficada

Elaborado por: Investigador

6. Si se desea reproducir la señal introducida, nos dirigimos al botón **Reproducir**, el cual nos reproduce la señal de audio ya filtrada. Figura 6.19.



Figura 6.19 Botón Reproducir

Elaborado por: Investigador

7. Para transformar nuestra señal de voz introducida a texto, hacemos un clic en el botón **Transformar**, el cual nos despliega las transformadas de las señales introducidas y las de nuestra base de datos, guardadas previamente. Figura 6.20

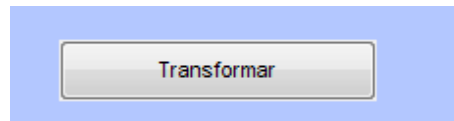
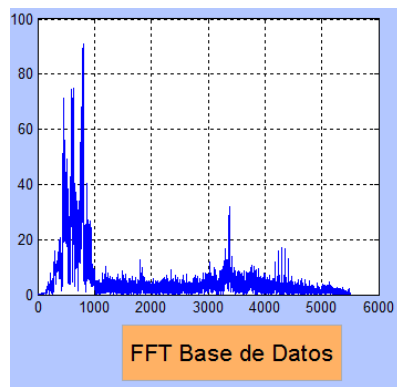


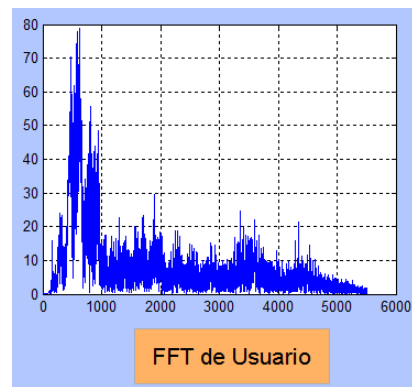
Figura 6.20 Botón Transformar

Elaborado por: Investigador

A continuación se muestra los gráficos de las transformadas después de pulsar el botón **Transformar**. Figura 6.21 y Figura 6.22.



a) Figuras 6.21



b) Figura 6.22

a) Transformada de la Señal de Base de Datos

b) Transformada de la Señal Introducida

Elaborado por: Investigador

Éstas gráficas son presentadas, para tener claro el proceso de transformación y comparación entre la señal introducida y la señal guardada en la base de datos.

8. A continuación podremos observar en la parte inferior de la ventana de la aplicación en forma textual nuestra señal introducida. Figura 6.23.



Figura 6.23 Letra Reconocida

Elaborado por: Investigador

En la Figura 6.24, se presenta de manera general como queda expresado nuestro programa después de haber realizado el proceso de transformación de una señal a texto.

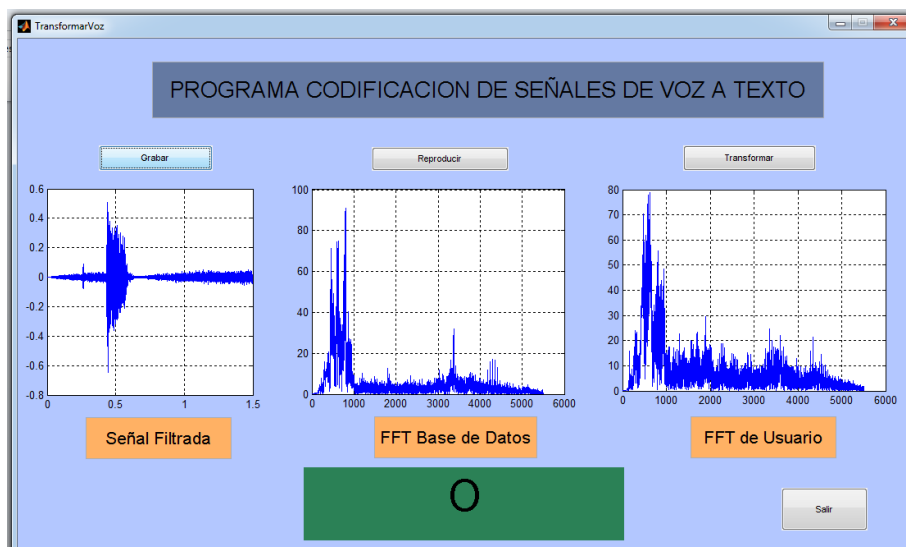


Figura 6.24. Pantalla Principal después del Proceso de Transformación  
Elaborado por: Investigador

9. Finalmente si deseamos salir de nuestro programa, nos dirigimos al botón **Salir** que se encuentra en la parte inferior derecha de nuestra aplicación. Figura 6.25.



Figura 6.25. Botón Salir  
Elaborado por: Investigador



### 6.7.7. Ejemplos de Transformación de Señales Voz a Texto

Una vez realizada la guía de uso, se puede realizar una corta demostración sobre el funcionamiento del programa con algunos ejemplos explicados a continuación.

- **Transformación a una Palabra:**

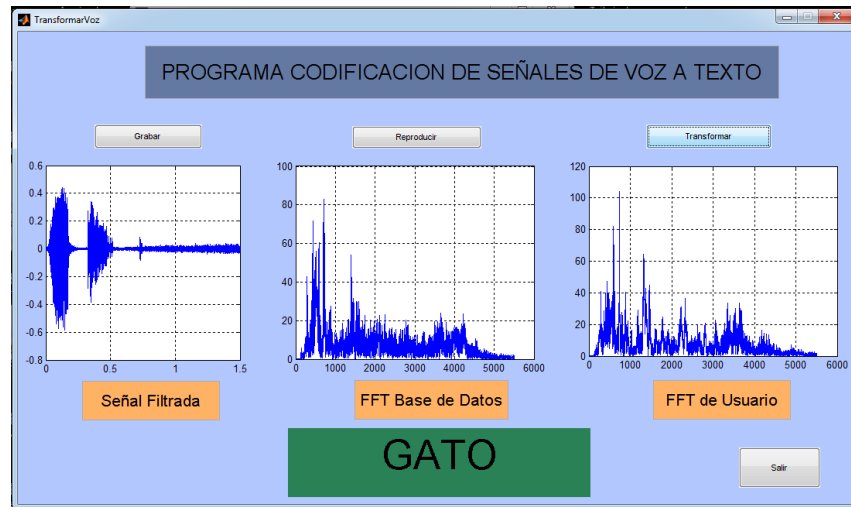


Figura 6.26. Ejemplo de transformación a una palabra

Elaborado por: Investigador

En la figura anterior se muestra como el programa transforma o reconoce una palabra expresada por el usuario. Las palabras que se desea transformar deben ser expresadas con total claridad para evitar una confusión en el programa, y por tanto no muestre la palabra pronunciada.

- **Transformación a un Número:**

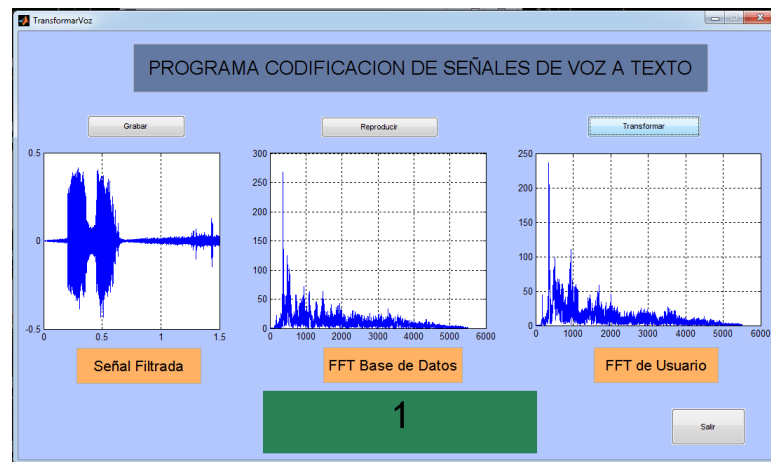


Figura 6.27. Ejemplo de transformación a un número

Elaborado por: Investigador

En la figura anterior se tiene un ejemplo de transformación o reconocimiento de un número expresado por el usuario. Los números expresados oralmente, son las palabras con mayor tendencia a sufrir confusión en el programa, por el parecido en la pronunciación con otros fonemas, ya sea de vocales, letras o palabras.

- **Transformación a una Vocal:**

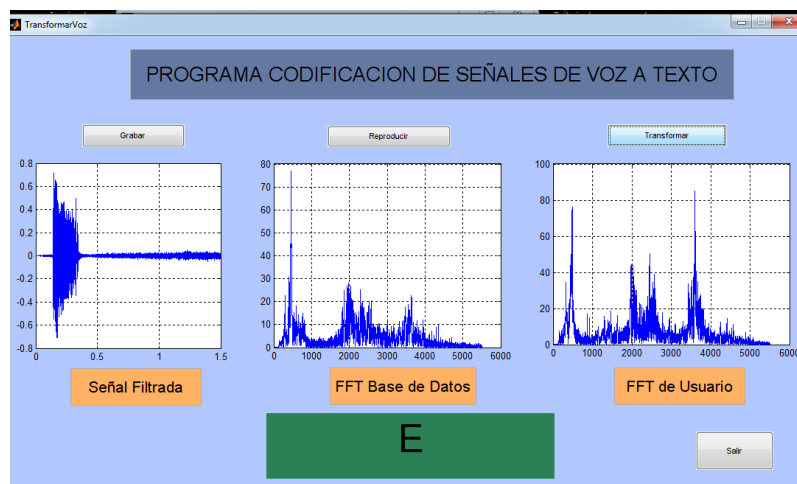


Figura 6.28. Ejemplo de transformación a una vocal

Elaborado por: Investigador

En la figura anterior se muestra la transformación de una señal de voz a vocal. Las transformaciones realizadas en el presente programa se encuentran en el anexo C.

### **6.7.8. Análisis e Interpretación de Resultados**

#### **6.7.8.1. Resultados de la Evaluación en Ambientes Ideales**

Se tomó señales de 5 personas de distintas edades, pero en un ambiente libre de ruidos externos, con el fin de obtener una señal más pura. A continuación se presenta una tabla con los datos adquiridos. Hay que resaltar que dicho análisis se realizó en base a las vocales.

Tabla 6.1. Tabla de Resultados en Ambientes Ideales

Elaborado por: Investigador

PERSONA	EDAD	Vocales					BANDA FRECUENCIA	MIN. ERROR	P. ERROR
		a	e	i	o	u			
A	20	OK					400 - 1000	5.1742	6.4594
			OK				200 a 600; 1800 a 2200	3.0073	5.1295
				OK			200 a 400; 3000 a 3700	4.8446	6.5961
					OK		400 - 900	6.1520	7.4241
						OK	300 - 700	4.5146	5.9489
B	16	OK					500 - 900	4.9654	5.9918
			OK				300 a 700; 2000 a 2500	5.3878	7.4479
				OK			100 a 500; 3200 a 3600	6.0041	7.0362
					OK		400 - 1000	4.6645	6.5291
						OK	250 - 600	5.6145	6.2134
C	32	OK					600 - 1100	5.082	6.2555
			OK				250 a 500; 1300 a 1700	5.9524	7.3777
				OK			100 a 350; 3000 a 3500	6.8557	8.4152
					OK		350 - 900	6.9687	7.9808
						OK	250 - 800	8.0446	8.9482
D MUJER	22	OK					800 - 1300	5.9147	6.8532
			OK				500 a 800; 2600 a 3150	6.2746	7.8463
				OK			200 a 550; 2900 a 3900	6.7654	7.8973
					OK		500 - 900	5.1343	6.6927
						OK	550 - 850	4.5978	6.6095
E		OK					600 - 1200	5.2635	6.5975
			OK				200 a 500; 1900 a 2300	3.2563	5.3145
				OK			350 a 750; 3100 a 3600	6.9124	8.5319
					OK		300 - 750	4.6018	5.9686
						OK	250 - 600	5.6389	6.2569

Tabla 6.2. Tabla de Promedio de Resultados en Ambientes Ideales

Elaborado por: Investigador

Promedio	OK					580 - 1100	4.3652	6.4315
		OK				290 a 620; 1560 a 2370	4.7757	6.6232
			OK			190 a 510; 3040 a 3660	6.2764	7.6953
				OK		390 - 890	5.5043	6.9191
					OK	320 - 710	5.6821	6.7954

En las pruebas realizadas en ambientes ideales, se pudo realizar la transformación de las señales de voz a texto de manera instantánea y eficaz, por el motivo de que no existía ruido externo, el cual nos puede producir señales no deseadas con las cuales el programa pueda tener fallas al momento de procesar la información. Por tal razón se establece un nivel de confiabilidad del 90% y 100%.

### 6.7.8.2. Resultados de la Evaluación en Ambientes con Ruido Moderado

Tabla 6.3 Tabla de Resultados en Ambientes con Ruido Moderado

Elaborado por: Investigador

PERSONA	EDAD	Vocales					BANDA FRECUENCIA	MIN. ERROR	P. ERROR
		a	e	i	o	u			
A	20	OK					600 - 1200	8.2513	9.3199
			OK				200 a 500; 1500 a 2000	7.5571	9.2749
				OK			400 a 800; 3100 a 3900	7.4421	8.1975
					OK		500 - 900	11.8382	12.7545
						OK	300 - 850	8.5146	9.5925
B	16	OK					450 -1000	6.4085	7.3496
			OK				400 a 800; 1500 a 1900	10.1171	10.8111
				OK			200 a 500; 3100 a 3500	8.6228	9.1435
					OK		400 - 900	7.2266	8.7426
						OK	300 - 750	9.4006	10.3970
C	32	OK					500 - 1150	6.6832	7.5951
			OK				450 a 850; 1400 a 2000	8.9524	9.7145
				OK			300 a 700; 2900 a 3450	8.4531	9.6348
					OK		300 - 800	9.2157	10.5912
						OK	350 - 800	9.1536	10.7135
D MUJER	22	OK					650 - 1400	9.3512	10.5276
			OK				600 a 950; 2800 a 3450	10.7235	11.3251
				OK			500 a 850; 3200 a 3700	8.9536	9.4452
					OK		450 - 900	10.3215	11.5213
						OK	500 - 950	9.6847	10.7856
E	18	OK					450 - 1000	7.2365	8.6325
			OK				400 a 700; 1600 a 2100	9.1467	10.0135
				OK			300 a 800; 3000 a 3650	8.9985	10.1537
					OK		300 - 800	9.3178	10.4235
						OK	400 - 900	9.7125	11.0475

Tabla 6.4. Tabla de Promedio de Resultados en Ambientes con Ruido Moderado

Elaborado por: Investigador

TOTAL	OK				530 - 1150	7.5861	8.6849
		OK			249 a 750; 1500 a 2289	9.2994	10.2278
			OK		340 a 730; 3200 a 3640	8.4940	9.3149
				OK	390 - 860	9.5840	10.8066
					OK	370 - 850	9.2932

En las pruebas realizadas en ambientes con ruidos moderados, se pudo efectuar la transformación de las señales de voz a texto de manera normal, pero en fonemas como “a” y “u” con ruido externo el programa no pudo procesar con facilidad. No se realiza la comparación con la base de datos y el programa tiende a confundir. Por tal razón se establece un nivel de confiabilidad del 80% y 90%.

### 6.7.8.3. Resultados de la Evaluación con Alto Ruido.

En las pruebas realizadas en ambientes con alta incidencia de ruido, no se pudo efectuar la transformación de las señales de voz a texto con eficacia, y se genera más cantidad de problemas con los fonemas “a”, “o”, “e” y “u”, solo la vocal “i” tiene más facilidad de ser reconocida por el acento agudo en la pronunciación. Por tal motivo los comandos del programa tienden a confundir todas las señales de entrada. En consecuencia se establece un nivel de confiabilidad bajo, pero si se pronuncian las vocales a una distancia cercana del micrófono y elevando la voz, existirá una mayor probabilidad de que sean reconocidas por el programa.

El presente punto no posee tabla ya que la aplicación no realiza con normalidad la transformación de las señales de voz introducidas con alto ruido.

## CAPÍTULO VII

### 7. CONCLUSIONES Y RECOMENDACIONES.

#### 7.7. Conclusiones.

- Los procesos matemáticos utilizados como la Transformada Rápida de Fourier fueron herramientas útiles y necesarias para el desarrollo de la aplicación.
- Las codificaciones de las señales de voz a texto se realizan con normalidad, aún incluso cuando se lo trabaja en ambientes con ruido moderado.
- Los datos generados en el programa depende de la vocalización, tipo de persona y del acento utilizado en cada pronunciación.
- EL filtro empleado en la aplicación, trabaja de mejor manera si la señal introducida no contiene ruido.
- Los fonemas con mayor facilidad de reconocimiento utilizados en el estudio son “e”, “i”, “o”. Entonces el programa lo transforma con normalidad independientemente del tipo de locutor.

## **7.8. Recomendaciones.**

- Al momento de generar la base de datos, se sugiere que el locutor pronuncie claramente los fonemas o palabras.
- Cuando se realice la transformación de señales de voz, de preferencia hacerlo en un ambiente libre de ruido, con el fin de obtener una señal pura.
- Si la transformación de la señal de voz no se realiza correctamente, volver a intentar pronunciando de mejor manera.
- Realizar un plan piloto de enseñanza con nuestra aplicación, con el fin de estudiar la reacción de los estudiantes frente al aprendizaje de una nueva forma de comunicación.
- Al implementar la aplicación en el centro educativo, realizar un seguimiento al proceso enseñanza – aprendizaje a largo plazo, con el fin de evaluar la evolución de los estudiantes en el reconocimiento de los símbolos del alfabeto así como de los números.



## Bibliografía.

- Figueira Carlos, (1996). *Datos, Señales y Codificación*.
- Jayant N. S. Peter Noll. (1984). *Digital Coding of Waveforms*. Prentice Hall
- Karris Steven, (2008). *Signal and Systems with Matlab Computing and Simulink Modeling*. (4<sup>th</sup> Ed.). USA. Orchard Publications.
- Lara Aznar Belén, (2004). *Codificación de Datos – Nuevas Tecnologías en Comunicaciones Móviles*. Revista Digital “Investigación y Educación”.
- Lim J. S., (1990). *Advances Topics in Signals Processing* .Prentice Hall
- Mariño Acebal José B, (1995). *Tratamiento Digital de la Señal*. España. Grupo de Procesado de Señal.
- Proakis John, (1996). *Digital Signal Processing: Principles, Algorithms, and Applications*, (Third Ed.). New Jersey. Prentice Hall.
- Ruiz Mezcua Belén, (2005). *La Voz y su Espectro*, Revista Sigma 27.
- Robles Moreno Alba, Discapitados Sensoriales
- Stranneby Dag, (2001). *Digital Signal Processing: DSP and Applications*. London. Florence Production Ltd.
- COUCH, Leon. (1998). *Sistemas de Comunicación Digitales y Analógicos* (5 ed.). México: Prentice Hall.
- García de Jalón, 2005, *Aprenda Matlab 7.0*, Madrid

## Enlaces:

- [www.blogspot.com](http://www.blogspot.com)  
Recuperada 1 octubre 2009 desde:  
<http://lonely113.blogspot.com> – Procesamiento Digital de Señales con Matlab.
- [www.mathworks.com](http://www.mathworks.com)  
Recuperada 17, noviembre 2010  
[www.mathworks.com/products/signal/Signal-Processing -Toolbox-6.15](http://www.mathworks.com/products/signal/Signal-Processing-Toolbox-6.15)
- [Kioskea.net](http://www.kioskea.net)  
Recuperada en octubre 16, 2008 desde:  
<http://es.kioskea.net/contents/transmission/transintro.php3>
- [www.sonidoyaudio.com](http://www.sonidoyaudio.com)  
Recuperada en  
[http://www.sonidoyaudio.com/sya/vp-tid:2-pid:19tipos\\_de\\_microfonos.html](http://www.sonidoyaudio.com/sya/vp-tid:2-pid:19tipos_de_microfonos.html) [www.SonidoyAudio.com](http://www.SonidoyAudio.com)
- [www.cinu.org](http://www.cinu.org)  
Recuperada en  
[www.cinu.org](http://www.cinu.org)  
[http://www.cinu.org.mx/temas/desarrollo/dessocial/integracion/p\\_dis.htm](http://www.cinu.org.mx/temas/desarrollo/dessocial/integracion/p_dis.htm)
- [www.un.org](http://www.un.org)  
Recuperada en  
<http://www.un.org/spanish/esa/social/disabled/dis50y10.htm>

# **ANEXOS**

## ANEXO A

### ENCUESTA.

La encuesta está dirigida a la población que conforma los docentes del Instituto de Educación Especial Dr. Camilo Gallegos, indicada con preguntas de carácter general y fáciles de entender, por lo tanto también fáciles de responder.

El esquema utilizado es el siguiente:

**UNIVERSIDAD TECNICA DE AMBATO**

**F.I.S.E.I.**

**ELECTRÓNICA Y COMUNICACIONES**

**Objetivo:** Con el fin de evaluar la factibilidad de la realización de un sistema de transformación de señales de voz a texto, usted tendrá que responder el siguiente cuestionario.

**Nota:** Lea detenidamente cada pregunta y marque su respuesta con una X.

1. ¿Cuál es el medio que utiliza para comunicarse con los alumnos con discapacidad auditiva?

Lenguaje de Señas ( )      Lectura de Labios ( )

2. ¿Conoce usted si la discapacidad auditiva de los alumnos es degenerativa?

SI ( )      NO ( )

3. ¿Conoce usted si existen alumnos que antes podían escuchar?

SI ( )      NO ( )

4. ¿Cree usted que con la ayuda de la tecnología, los alumnos sordo-mudos puedan tener la oportunidad de desarrollar nuevos métodos de comunicación?

SI ( )      NO ( )

5. ¿Los alumnos conocen los símbolos alfanuméricos?

SI ( )      NO ( )

6. ¿Cree usted que con la implementación del software los alumnos podrán desenvolverse de mejor manera en la sociedad?

SI ( )      NO ( )

7. ¿Piensa usted que en un futuro las personas con deficiencia auditiva puedan desarrollar la habilidad de leer textos en español?

SI ( )      NO ( )

¡Gracias por su colaboración!

## ANEXO B

### Código del programa “Transformar Voz a Texto”

```
function varargout = TransformarVoz(varargin)

% Codigo Inicializado – NO EDITAR

gui_Singleton = 1;
gui_State = struct('gui_Name',    mfilename, ...
                  'gui_Singleton', gui_Singleton, ...
                  'gui_OpeningFcn', @TransformarVoz_OpeningFcn, ...
                  'gui_OutputFcn', @TransformarVoz_OutputFcn, ...
                  'gui_LayoutFcn', [] , ...
                  'gui_Callback', []);
if nargin && ischar(varargin{1})
    gui_State.gui_Callback = str2func(varargin{1});
end

if nargout
    [varargout{1:nargout}] = gui_mainfcn(gui_State, varargin{:});
else
    gui_mainfcn(gui_State, varargin{:});
end

% Fin del código de inicialización- NO EDITAR

%---- Ejecuta justo antes TransformarVoz Y se hace visible

function TransformarVoz_OpeningFcn(hObject, eventdata, handles, varargin)

handles.output = hObject;

guidata(hObject, handles);

%---- Función de salida del programa Compresor

function varargout = TransformarVoz_OutputFcn(hObject, eventdata, handles)

varargout{1} = handles.output;
```

```
% --- Ejecuta durante la creación del objeto, después de establecer todas las propiedades
```

```
function edit1_CreateFcn(hObject, eventdata, handles)
```

```
if ispc && isequal(get(hObject,'BackgroundColor'),
```

```
get(0,'defaultUicontrolBackgroundColor'))
```

```
set(hObject,'BackgroundColor','white');
```

```
end
```

```
--- BOTON GRABAR ---
```

```
function BotonGrabar_Callback(hObject, eventdata, handles)
```

```
clc %limpia la pantalla
```

```
global y fs; %Establecer variables globales
```

```
fs=11025; %frecuencia de muestreo
```

```
tiempograb=1.5; %Tiempo de grabacion
```

```
y=wavrecord(tiempograb*fs,fs,1); %función de grabacion
```

```
soundsc(y,fs); %Reproduce grabación
```

```
ts=1/fs;
```

```
t=0:ts:tiempograb-ts;
```

```
b=[1 -0.95];
```

```
yf=filter(b,1,y); %Proceso de filtrado
```

```
len = length(y); %longitud del vector
```

```
avg_e = sum(y.*y)/len; %promedio señal entera
```

```
THRES = 0.2;
```

```
soundsc(y,fs) %Reproduce señal filtrada
```

```
wavwrite(yf,fs,'voz'); %Graba en archivo .wav
```

```
%---- Grafica señal grabada
```

```
figure(1);
```

```
plot(t,y);grid on;
```

```
set(handles.axes1); % Establece los ejes de graficación
```

```
axes(handles.axes1);
```

```
plot(t,yf);grid on; % Grafica en los axes
```

```
%----- Mensaje Fin de Grabacion
```

```
msgbox('Grabación Terminada');  
guidata(hObject, handles);
```

```
--- BOTON REPRODUCIR ---
```

```
function BotonReproducir_Callback(hObject, eventdata, handles)
```

```
[y,fs]=wavread('voz'); %Lectura del archivo grabado
```

```
soundsc(y,fs); %Reproducción de archivo grabado
```

```
----- Función Normalizar -----
```

```
% Función que será utilizada en los siguientes procesos
```

```
function sonidoN=normalizar(sonido)
```

```
maximo=max(abs(sonido));
```

```
n=length(sonido); %calcula el tamaño del vector
```

```
sonidoN=zeros(1,n);
```

```
for i=1:1:n
```

```
    sonido(i)=sonido(i)/maximo;
```

```
end
```

```
--- BOTON TRANSFORMAR ---
```

```
function BotonTransformar_Callback(hObject, eventdata, handles)
```

```
N=length(y); %calcula el tamaño del vector
```

```
f=(0:N-1)*fs/N;
```



```
% Parámetros entrenados para reconocer
```

```
v1 = wavread('a');  
v2 = wavread('e');  
v3 = wavread('i');  
v4 = wavread('o');  
v5 = wavread('u');
```

```
p1 = wavread('b');  
p2 = wavread('c');  
p3 = wavread('d');  
p4 = wavread('f');  
p5 = wavread('g');  
p6 = wavread('h');  
p7 = wavread('j');  
p8 = wavread('k');  
p9 = wavread('l');  
p10= wavread('m');  
p11= wavread('n');  
p12= wavread('p');  
p13= wavread('q');  
p14= wavread('r');  
p15= wavread('s');  
p16= wavread('t');  
p17= wavread('v');  
p18= wavread('x');  
p19= wavread('y');  
p20= wavread('z');
```

```
n0= wavread('0');  
n1= wavread('1');  
n2= wavread('2');  
n3= wavread('3');  
n4= wavread('4');  
n5= wavread('5');  
n6= wavread('6');  
n7= wavread('7');  
n8= wavread('8');  
n9= wavread('9');
```

```
%--- Uso de la Función Normalizar
```

```
vocala = normalizar(v1);  
vocale = normalizar(v2);  
vocali = normalizar(v3);
```

```
vocalo = normalizar(v4);  
vocalu = normalizar(v5);
```

```
letra1 = normalizar(p1);  
letra2 = normalizar(p2);  
letra3 = normalizar(p3);  
letra4 = normalizar(p4);  
letra5 = normalizar(p5);  
letra6 = normalizar(p6);  
letra7 = normalizar(p7);  
letra8 = normalizar(p8);  
letra9 = normalizar(p9);  
letra10 = normalizar(p10);  
letra11 = normalizar(p11);  
letra12 = normalizar(p12);  
letra13 = normalizar(p13);  
letra14 = normalizar(p14);  
letra15 = normalizar(p15);  
letra16 = normalizar(p16);  
letra17 = normalizar(p17);  
letra18 = normalizar(p18);  
letra19 = normalizar(p19);  
letra20 = normalizar(p20);
```

```
numero0 = normalizar(n0);  
numero1 = normalizar(n1);  
numero2 = normalizar(n2);  
numero3 = normalizar(n3);  
numero4 = normalizar(n4);  
numero5 = normalizar(n5);  
numero6 = normalizar(n6);  
numero7 = normalizar(n7);  
numero8 = normalizar(n8);  
numero9 = normalizar(n9);
```

### %---- Transformada de las Letras y Números

```
transvocala = abs((fft(vocala)));  
transvocale = abs((fft(vocale)));  
transvocali = abs((fft(vocali)));  
transvocalo = abs((fft(vocalo)));  
transvocalu = abs((fft(vocalu)));
```

```
transletra1 = abs((fft(letra1)));  
transletra2 = abs((fft(letra2)));  
transletra3 = abs((fft(letra3)));
```

```
transletra4 = abs((fft(letra4)));
transletra5 = abs((fft(letra5)));
transletra6 = abs((fft(letra6)));
transletra7 = abs((fft(letra7)));
transletra8 = abs((fft(letra8)));
transletra9 = abs((fft(letra9)));
transletra10 = abs((fft(letra10)));
transletra11 = abs((fft(letra11)));
transletra12 = abs((fft(letra12)));
transletra13 = abs((fft(letra13)));
transletra14 = abs((fft(letra14)));
transletra15 = abs((fft(letra15)));
transletra16 = abs((fft(letra16)));
transletra17 = abs((fft(letra17)));
transletra18 = abs((fft(letra18)));
transletra19 = abs((fft(letra19)));
transletra20 = abs((fft(letra20)));
```

```
transnumero0 = abs((fft(numero0)));
transnumero1 = abs((fft(numero1)));
transnumero2 = abs((fft(numero2)));
transnumero3 = abs((fft(numero3)));
transnumero4 = abs((fft(numero4)));
transnumero5 = abs((fft(numero5)));
transnumero6 = abs((fft(numero6)));
transnumero7 = abs((fft(numero7)));
transnumero8 = abs((fft(numero8)));
transnumero9 = abs((fft(numero9)));
```

%----- Proceso de Reconocimiento

```
voz_usuario=wavread('voz');
```

```
usuario=normalizar(voz_usuario);
```

```
transusuario=abs((fft(usuario)));
```

```
error(1)=mean(abs(transvocala-transusuario));
error(2)=mean(abs(transvocale-transusuario));
error(3)=mean(abs(transvocali-transusuario));
error(4)=mean(abs(transvocalo-transusuario));
error(5)=mean(abs(transvocalu-transusuario));
error(6)=mean(abs(transletra1-transusuario));
error(7)=mean(abs(transletra2-transusuario));
error(8)=mean(abs(transletra3-transusuario));
error(9)=mean(abs(transletra4-transusuario));
error(10)=mean(abs(transletra5-transusuario));
```

```

error(11)=mean(abs(transletra6-transusuario));
error(12)=mean(abs(transletra7-transusuario));
error(13)=mean(abs(transletra8-transusuario));
error(14)=mean(abs(transletra9-transusuario));
error(15)=mean(abs(transletra10-transusuario));
error(16)=mean(abs(transletra11-transusuario));
error(17)=mean(abs(transletra12-transusuario));
error(18)=mean(abs(transletra13-transusuario));
error(19)=mean(abs(transletra14-transusuario));
error(20)=mean(abs(transletra15-transusuario));
error(21)=mean(abs(transletra16-transusuario));
error(22)=mean(abs(transletra17-transusuario));
error(23)=mean(abs(transletra18-transusuario));
error(24)=mean(abs(transletra19-transusuario));
error(25)=mean(abs(transletra20-transusuario));
error(26)=mean(abs(transnumero0-transusuario));
error(27)=mean(abs(transnumero1-transusuario));
error(28)=mean(abs(transnumero2-transusuario));
error(27)=mean(abs(transnumero3-transusuario));
error(28)=mean(abs(transnumero4-transusuario));
error(29)=mean(abs(transnumero5-transusuario));
error(30)=mean(abs(transnumero6-transusuario));
error(31)=mean(abs(transnumero7-transusuario));
error(32)=mean(abs(transnumero8-transusuario));
error(33)=mean(abs(transnumero9-transusuario));

```

```

min_error=min(error); %Almacena el menor error

```

```

display(error) %Muestra el vector error
display(min_error) %Muestra el menor error

```

---- Proceso de Comparaciones para identificar palabras ----

```

if(min_error == error(1))

```

```

    set(handles.axes2); %Establece el axes para graficar
    axes(handles.axes2); %Axes habilitado para graficar
    plot(f(1:N/2),transvocala(1:N/2));grid on; % Grafica del espectro
    de la letra de la base de datos

```

```

    set (handles.text1, 'string', 'A'); % Muestra la letra comparada

```

```

set(handles.axes3); %Establece el axes para graficar
axes(handles.axes3); %Axes habilitado para graficar

plot(f(1:N/2),transusuario(1:N/2));grid on; %Grafica del espectro
de la letra pronunciada

end

if(min_error == error(2))

    set(handles.axes2);
    axes(handles.axes2);
    plot(f(1:N/2),transvocale(1:N/2));grid on;

    set (handles.text1, 'string', 'E');

    set(handles.axes3);
    axes(handles.axes3);

    plot(f(1:N/2),transusuario(1:N/2));grid on;

end

if(min_error == error(3))

    set(handles.axes2);
    axes(handles.axes2);
    plot(f(1:N/2),transvocali(1:N/2));grid on;

    %msgbox('Usted dijo: I');
    set (handles.text1, 'string', 'I');

    set(handles.axes3);
    axes(handles.axes3);
    plot(f(1:N/2),transusuario(1:N/2));grid on;

end

if(min_error == error(4))

    set(handles.axes2);
    axes(handles.axes2);
    plot(f(1:N/2),transvocalo(1:N/2));grid on;

    set (handles.text1, 'string', 'O');

```

```
set(handles.axes3);  
axes(handles.axes3);  
plot(f(1:N/2),transusuario(1:N/2));grid on;
```

end

```
if(min_error == error(5))
```

```
set(handles.axes2);  
axes(handles.axes2);  
plot(f(1:N/2),transvocalu(1:N/2));grid on;
```

```
set (handles.text1, 'string', 'U');
```

```
set(handles.axes3);  
axes(handles.axes3);  
plot(f(1:N/2),transusuario(1:N/2));grid on;
```

end

```
if(min_error == error(6))
```

```
set(handles.axes2);  
axes(handles.axes2);  
plot(f(1:N/2),transpalabra1(1:N/2));grid on;
```

```
set (handles.text1, 'string', 'B');
```

```
set(handles.axes3);  
axes(handles.axes3);  
plot(f(1:N/2),transusuario(1:N/2));grid on;
```

end

```
if(min_error == error(7))
```

```
set(handles.axes2);  
axes(handles.axes2);  
plot(f(1:N/2),transpalabra2(1:N/2));grid on;
```

```
set (handles.text1, 'string', 'C');
```

```
set(handles.axes3);  
axes(handles.axes3);  
plot(f(1:N/2),transusuario(1:N/2));grid on;
```

```
end
```

Y se repite el proceso de comparación para todas las letras y números creados en la base de datos. Si se desea que la aplicación reconozca más palabras solo se tiene que grabar en la base de datos.

--- Botón Salir ----

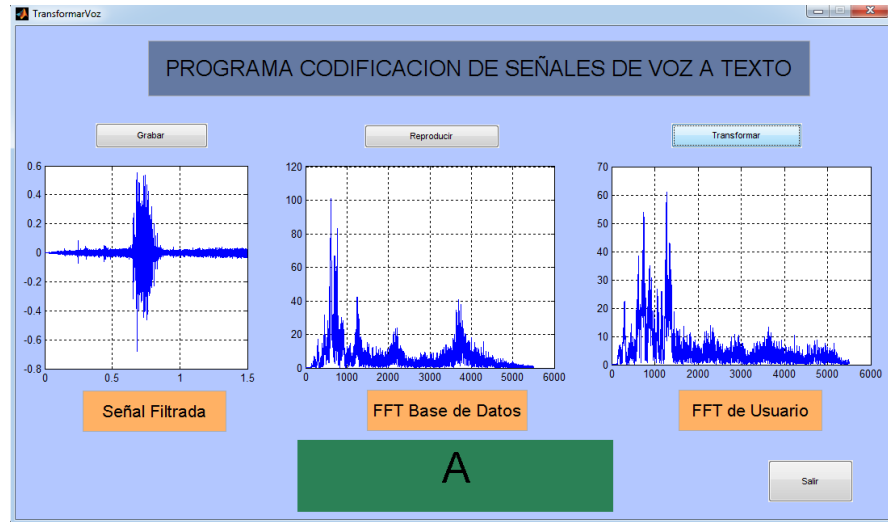
```
function BotonSalir_Callback(hObject, eventdata, handles)
```

```
exit; % Comando para salir de la aplicación
```

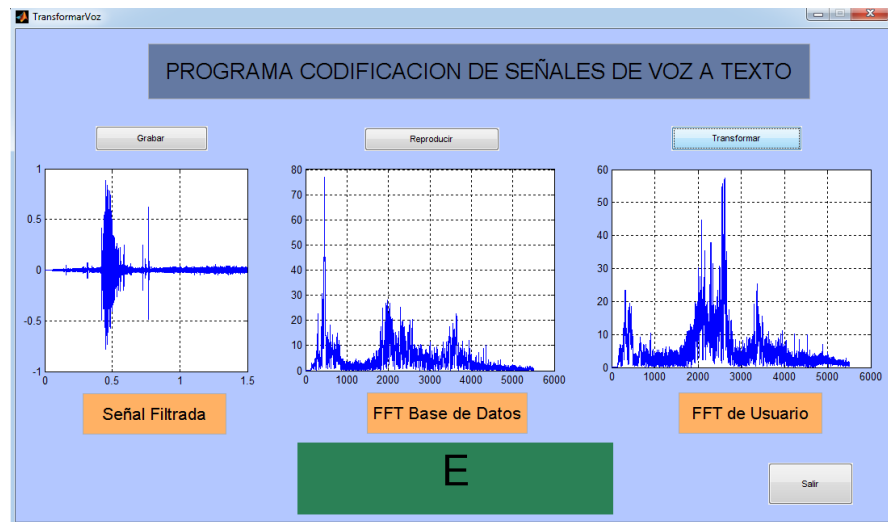
## ANEXO C

### CAPTURAS DE LAS VOCALES RECONOCIDAS

- Vocal “a”:

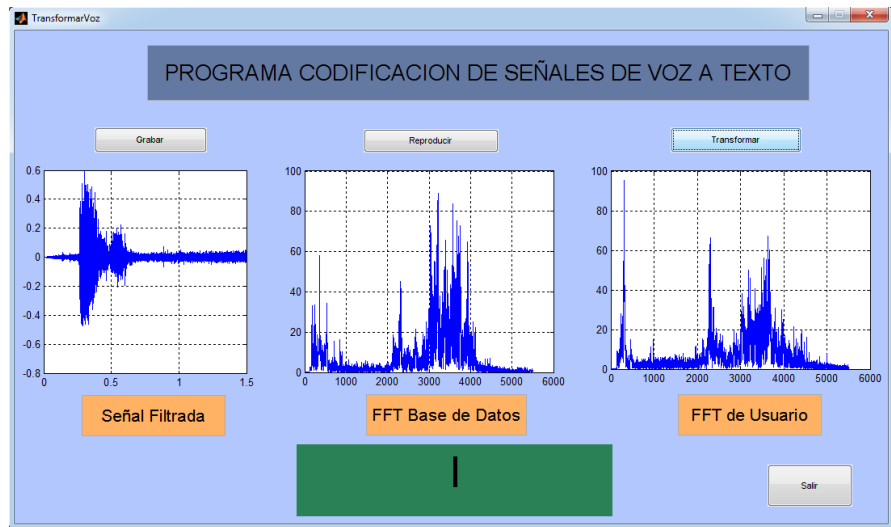


- Vocal “e”:

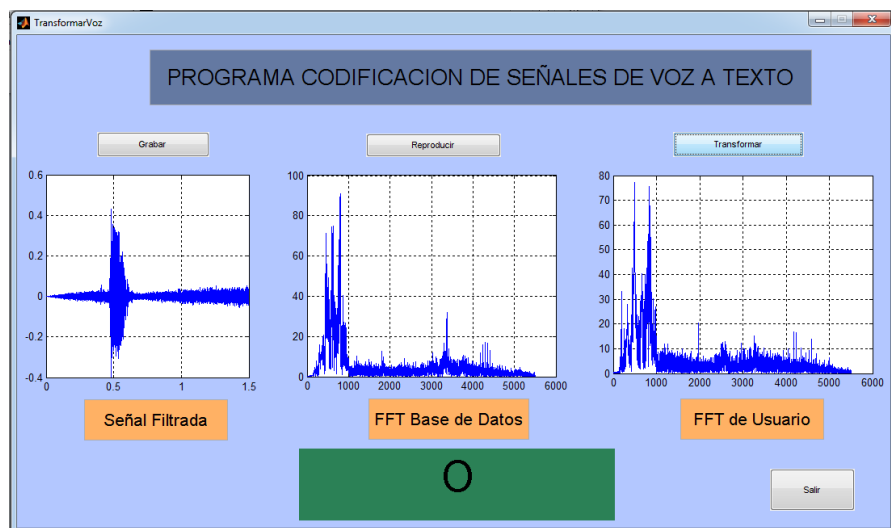




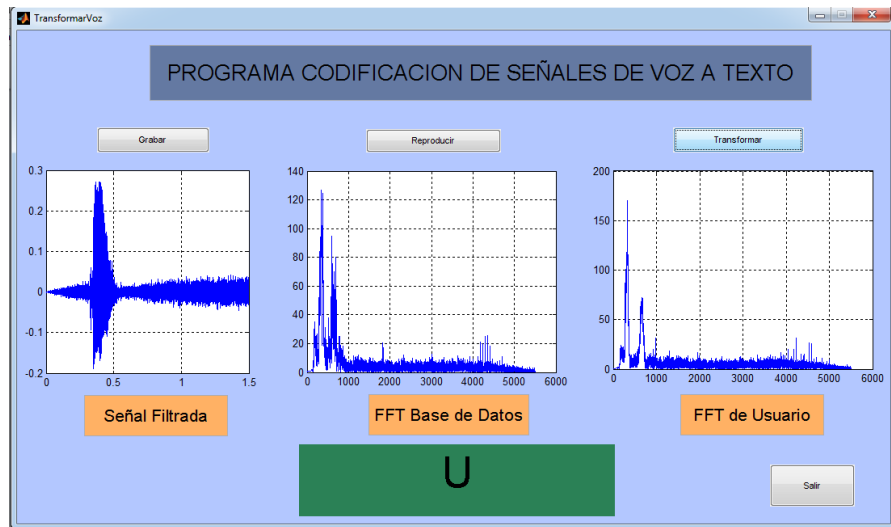
- Vocal “i”:



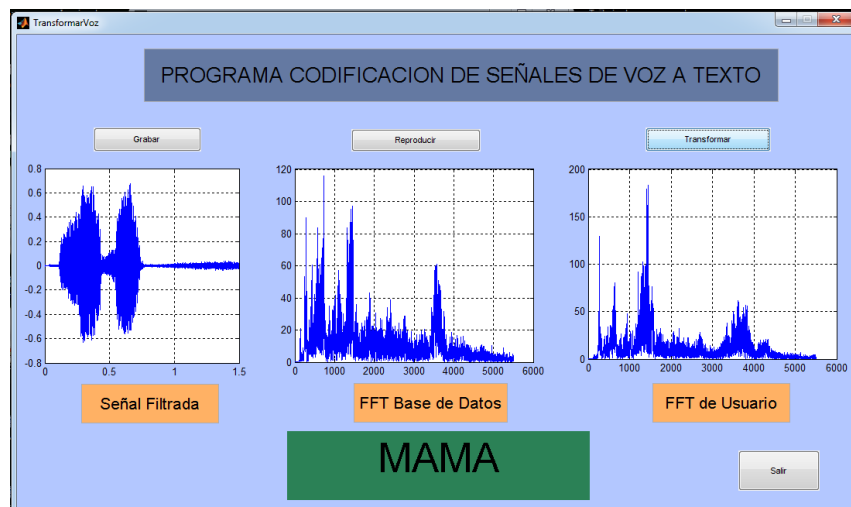
- Vocal “o”:



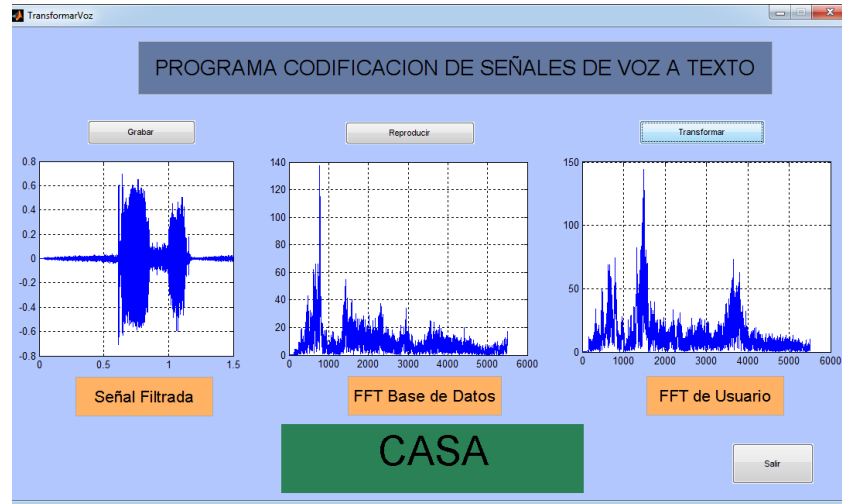
- Vocal “u”:



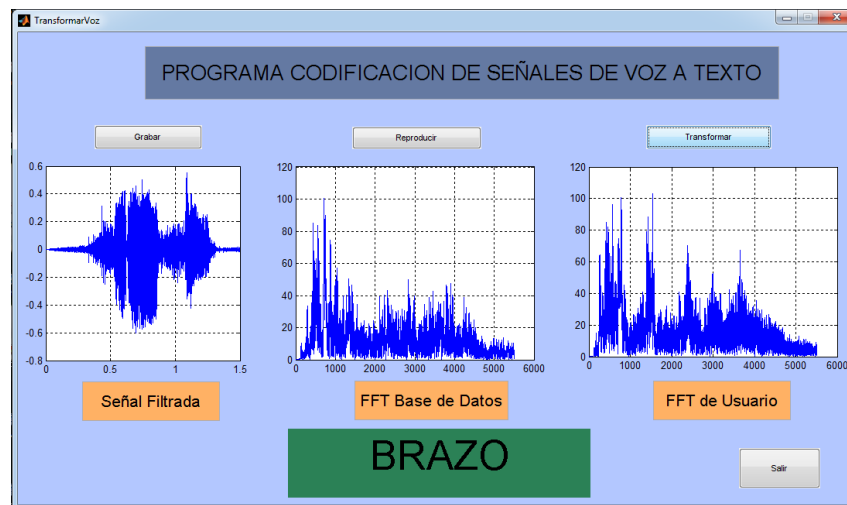
- Palabra



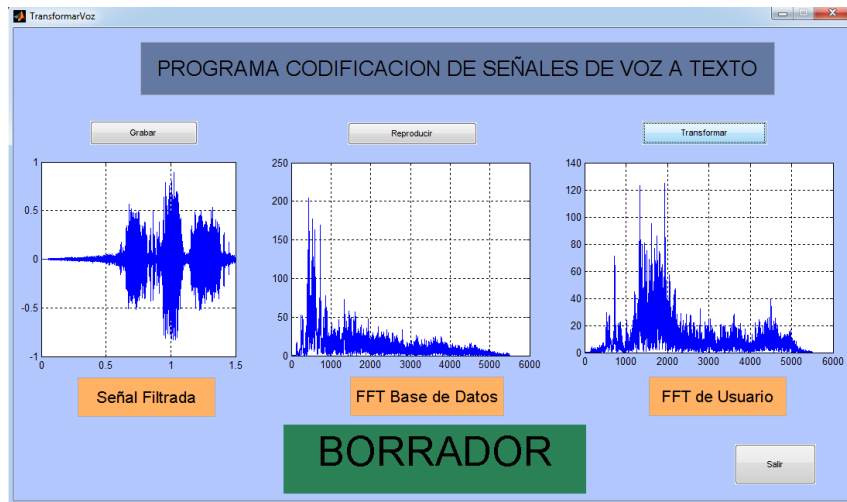
- Palabra



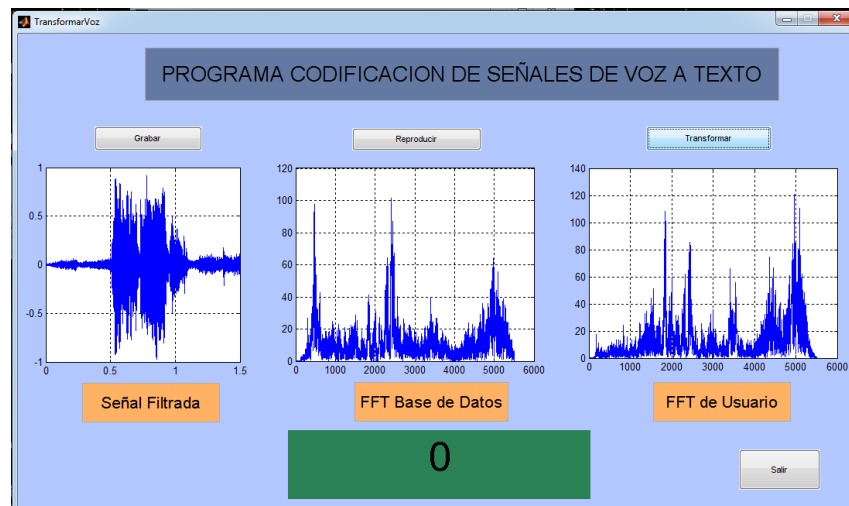
- Palabra



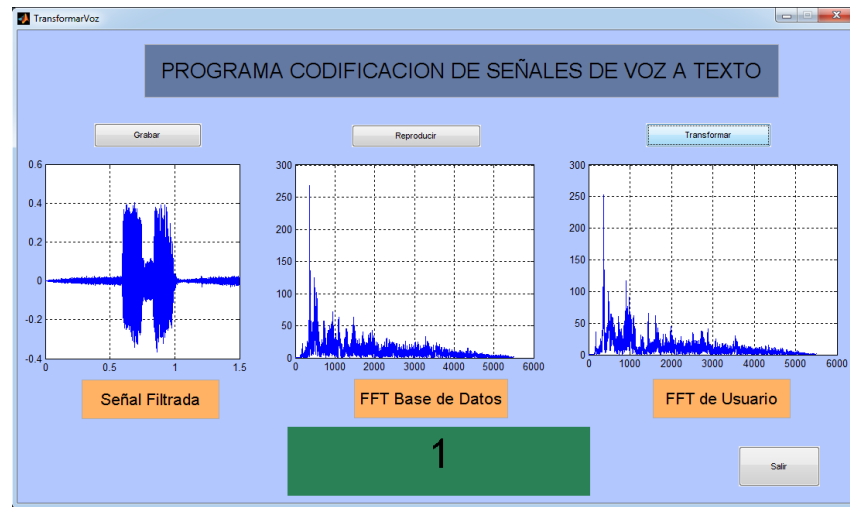
- Palabra



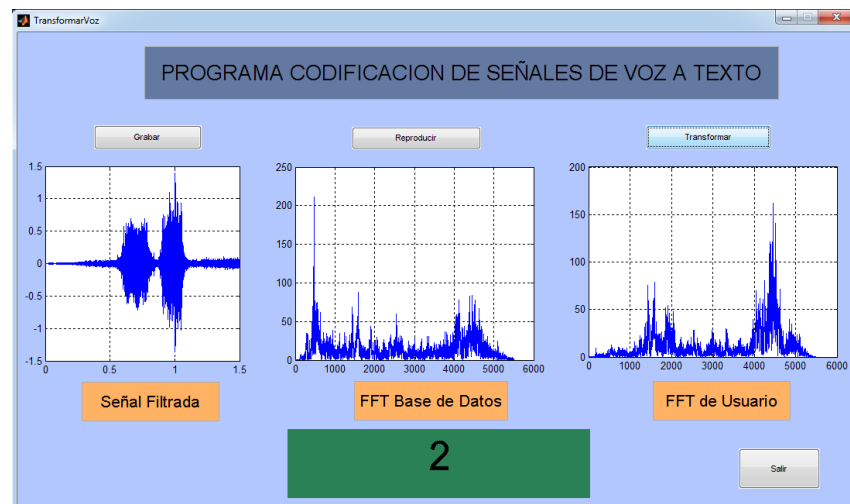
- Número "0"



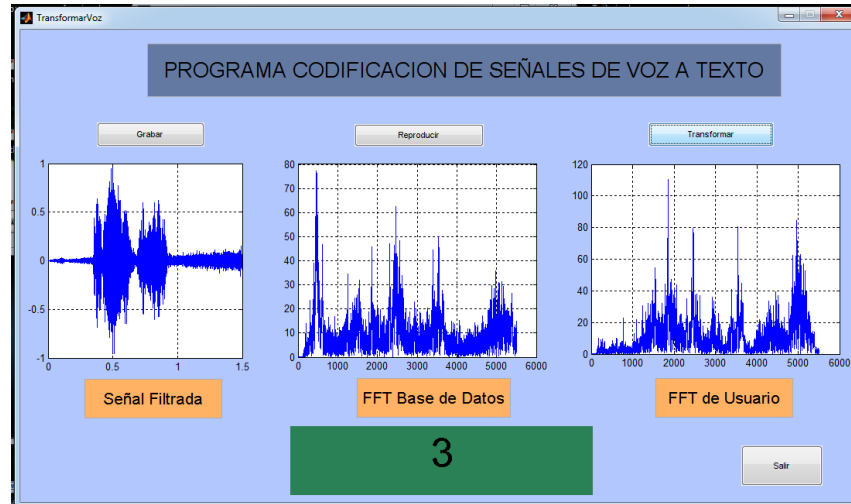
- Número “1”



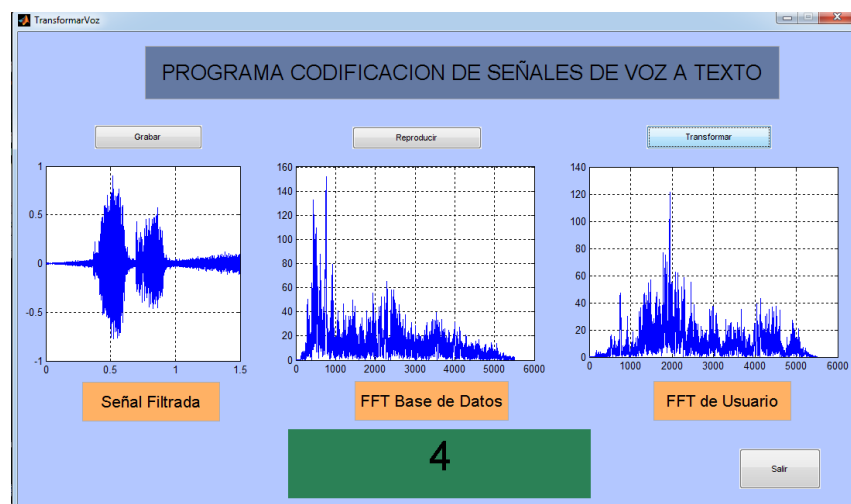
- Número “2”



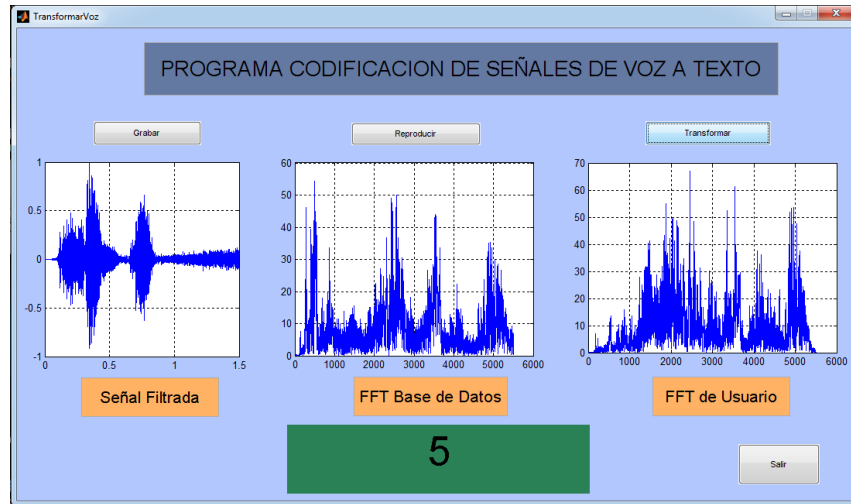
- Número “3”



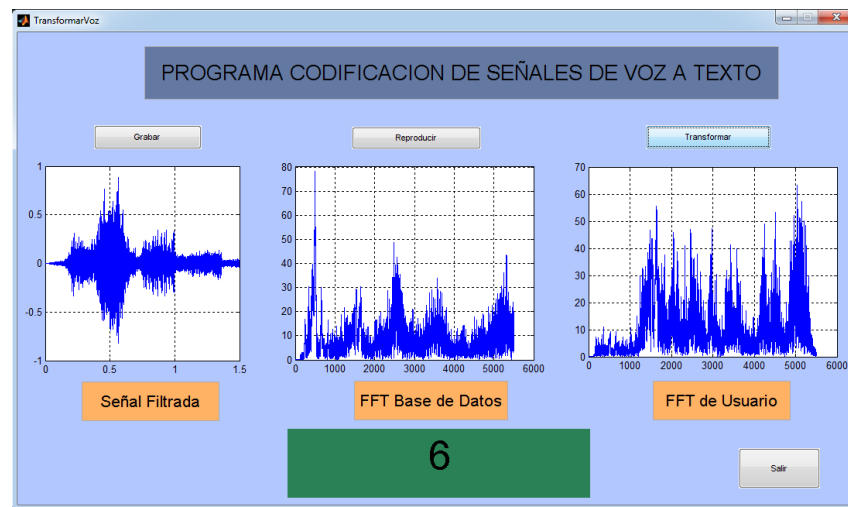
- Número “4”



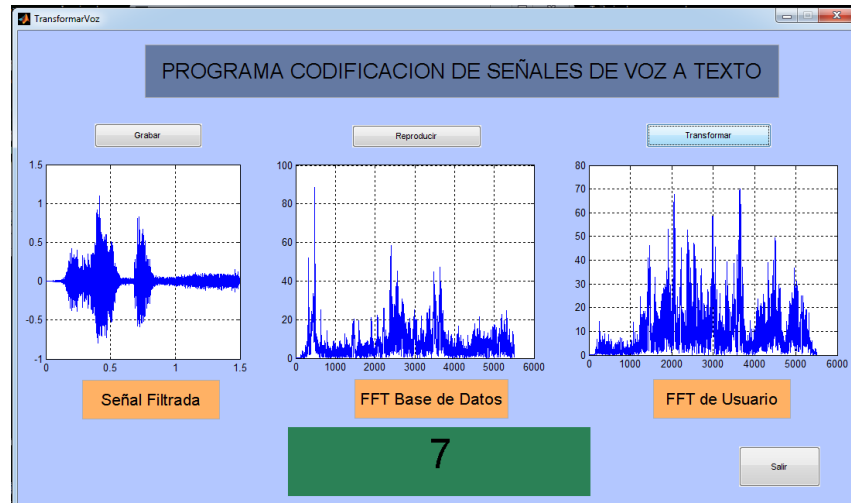
- Número “5”



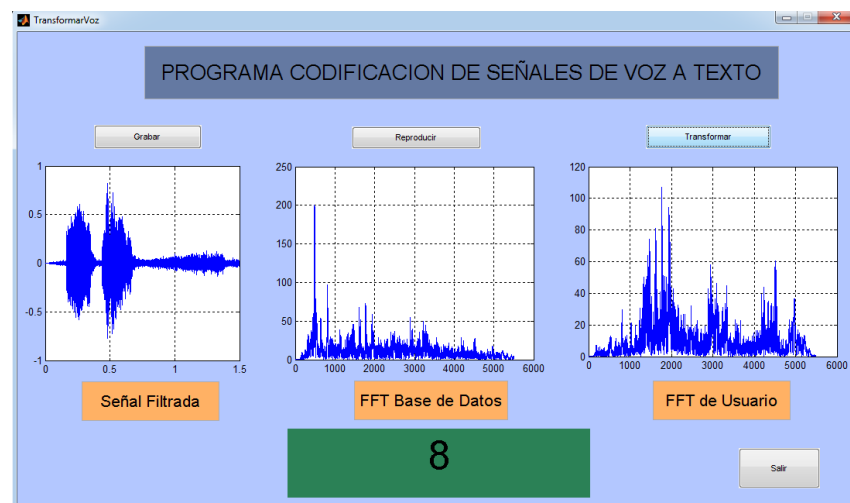
- Número “6”



- Número “7”

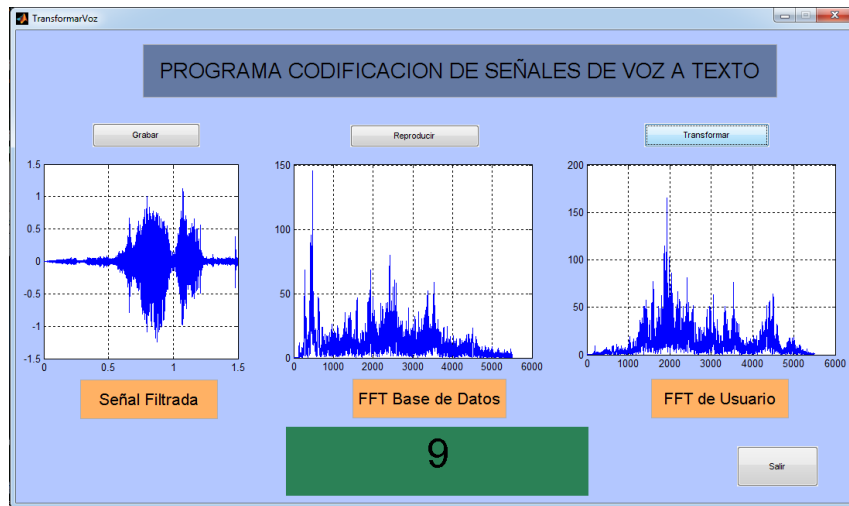


- Número “8”





- Número “9”



- Número “10”

